PATRICK SUPPES

INTRODUCCION A LA LOGICA SIMBOLICA

3a. IMPRESION

C. E. C. S. A.

INTRODUCCION
A LA
LOGICA SIMBOLICA

Por

PATRICK SUPPES

Profesor de Filosofía
Universidad de Stanford

COMPAÑIA EDITORIAL CONTINENTAL, S. A.
MEXICO – ESPAÑA – ARGENTINA – CHILE

SUCURSALES, DEPOSITOS Y REPRESENTACIONES EN:

Bolivia – Brasil – Colombia – Costa Rica – Dominicana – Ecuador – El Salvador
Estados Unidos – Guatemala – Honduras – Nicaragua – Panamá – Paraguay – Perú
Portugal – Puerto Rico – Uruguay – Venezuela

Título original en inglés:
INTRODUCTION TO LOGIC

Traducido por:
GABRIEL AGUIRRE CARRASCO
Catedrático de la Universidad Autónoma de Puebla

Edición autorizada por:
D. VAN NOSTRAND COMPANY, INC. – Princeton, New Jersey
Copyright © by D. Van Nostrand Company, Inc.
Library of Congress Catalog Card No. 57-8153

Primera edición en español: abril de 1966

Segunda impresión en español: marzo de 1969

Tercera impresión en español:
septiembre de 1970

Derechos Reservados © en Lengua Española–1966, Primera Publicación

COMPAÑIA EDITORIAL CONTINENTAL, S. A.
CALZADA DE TLALPAN NÚM. 4620, MÉXICO 22, D. F.

MIEMBRO DE LA CAMARA NACIONAL DE LA INDUSTRIA EDITORIAL
Registro Núm. 43

AV. REPÚBLICA ARGENTINA NÚM. 168, BARCELONA 6, ESPAÑA
ENTRE RÍOS NÚM. 1256, BUENOS AIRES, ARGENTINA
AMUNÁTEGUI NÚM. 458, SANTIAGO DE CHILE, CHILE

A la memoria de

J. C. C. McKINSEY

PREFACIO

Este libro ha sido escrito primordialmente para servir como libro de texto de un primer curso de lógica moderna. No supone una preparación previa en matemáticas o filosofía. Mi objetivo principal ha sido el de familiarizar al lector con una teoría exacta y completa de la inferencia lógica y mostrar cómo puede ser ésta usada en matemáticas y en las ciencias empíricas. Puesto que varios libros en circulación tienen fines estrechamente relacionados con el que se acaba de indicar, puede ser del caso mencionar los principales rasgos característicos del presente libro.

La Parte I (los primeros ocho capítulos) trata de los principios formales de la inferencia y la definición. Comenzando con la teoría de la inferencia oracional en el Cap. 2, hay un énfasis continuo en la aplicación del método de interpretación para demostrar la invalidez de los argumentos, la consistencia de las premisas, o los axiomas de una teoría independiente. Hay un intento detallado (Cap. 7) para relacionar la teoría formal de la inferencia con las pruebas informales típicas, comunes en toda la matemática. La teoría de la definición se presenta (Cap. 8) con más detalle que en cualquier otro texto conocido del autor; se incluye una discusión del método de Padoa para demostrar la independencia de conceptos primitivos.

La segunda parte (los cuatro últimos capítulos) está dedicada a la teoría elemental intuitiva de los conjuntos, con capítulos separados sobre conjuntos, relaciones y funciones. El tratamiento de las relaciones de ordenación en el Cap. 10 es un tanto extenso. La segunda parte es casi independiente y puede leerse independientemente de la primera parte. El último capítulo (Cap. 12) trata de la fundamentación teórica del método axiomático en la teoría de los conjuntos. La idea de que el mejor modo de fundamentar una rama de las matemáticas es definir adecuados predicados, en términos de la teoría de los conjuntos, es una idea familiar para los matemáticos modernos y, ciertamente, no es original del autor, pero la exposición de esta idea, que suministra una aguda fundamentación lógica del

método axiomático, ha sido omitida de los excelentes libros elementales de texto de matemática moderna que han aparecido en los últimos años.

Comenzando con el Cap. 4, se introducen en la discusión y ejercicios numerosos ejemplos de teorías formuladas axiomáticamente. Estos ejemplos van desde la teoría de los grupos y el álgebra de los números reales hasta la teoría elemental de la probabilidad, la mecánica clásica de puntos materiales y la teoría de la medición de intensidades de sensación. (El apartado sobre mecánica es la única excepción a la aseveración general de que no hay requisitos previos de índole matemática para la lectura de este libro; se requiere cierto conocimiento del cálculo diferencial e integral para la completa inteligencia de este apartado, que es el último del libro). Ciertos de los ejemplos incluidos en relación con estos ejemplos sustantivos, son más difíciles que los que generalmente se incluyen en un texto de lógica elemental. El propósito de estos ejercicios es el de estimular a los estudiantes más capaces. Hay, en cambio, un gran número de ejercicios adicionales de índole relativamente sencilla, que ilustran adecuadamente todos los principios generales establecidos. Se espera que el material sobre medición, probabilidad y mecánica del Cap. 12 pueda ser útil en algunos cursos de filosofía o de ciencia.

El sistema de inferencia para la lógica de predicados de primer grado desarrollado en los Caps. 2, 4 y 5, ha sido diseñado para corresponder tan estrechamente como sea posible al concepto del autor de las técnicas más naturales de prueba informal. Probablemente el aspecto más nuevo del sistema es el método de manejar cuantificadores existenciales mediante "nombres ambiguos"; la idea central de este método se relaciona con el símbolo ε de Hilbert. Puesto que muchos maestros de lógica tienen sus propias reglas preferidas para manejar inferencias con cuantificadores, debe mencionarse que las reglas introducidas juegan un papel importante sólo en los Caps. 4 y 5.

Numerosas personas han contribuido al desarrollo gradual de este libro. Estoy particularmente en deuda con el Prof. Robert McNaughton por muchas críticas y sugestiones útiles, basadas en su experiencia docente con anteriores borradores de este libro; al Prof. Herman Rubin, quien contribuyó a la formulación del sistema de deducción natural presentado en los Caps. 2 y 4 y al Sr. Dana Scott por muchas útiles sugestiones relativas a los Caps. 8 y 12. También estoy en deuda con el Prof. Moffatt Hancock, de la Stanford Law School por varios ejercicios de los Caps. 2 y 4. Varios profesores adjuntos de Stanford han ayudado en la preparación de ejercicios y han hecho numerosas observaciones útiles; muy especialmente el

Sr. Leonard Leving, la Sra. Muriel Wood Gerlach y la Sra. Rina Ullmann.

El último conjunto de revisiones se ha beneficiado de los comentarios y observaciones de los profesores Ernest Adams, Herman Chernoff, Benson Mates, John Myhill, David Nivison, Hartley Rogers, Jr., Leo Simons, Robert Vaught y el Sr. Richard Robinson. Las observaciones extraordinariamente detalladas y perspicaces del Prof. Vaught fueron especialmente valiosas. La Srita. Peggy Reis y la Sra. Karol Valpreda Walsh prestaron importante ayuda en la lectura de las pruebas. La Srita. Reis, además, ha mecanografiado gustosa y eficientemente las diversas ediciones preliminares usadas en clase desde el verano de 1954.

Este libro está dedicado a la memoria del Prof. J. C. C. McKinsey. Se planeó originalmente escribir un libro en colaboración con él, pero debido a su prematura muerte en 1953, resulta poco notorio ahora su competente trabajo. Escribió los borradores originales de los Caps. 9, 10 y 11, pero estos capítulos han sido revisados tres veces, y la longitud de los Caps. 10 y 11 se ha más que duplicado. Los otros nueve capítulos son de mi sola responsabilidad, en todos respectos.

PATRICK SUPPES

PRESENTACION DE LA EDICION EN ESPAÑOL

La lógica simbólica, juntamente con las matemáticas, son las dos ciencias estructurales de la civilización moderna y, si tenemos en cuenta que esta lógica es el fundamento de la matemática actual, podemos entonces decir que la lógica simbólica es la ciencia básica de la civilización de nuestros días, y la matriz de las formas del futuro.

En sus cien años de vida ha llegado a los circuitos lógicos de los cerebros electrónicos; este hecho es un espectacular indicio de su influjo.

La lógica aristotélica puede considerarse, de hecho, liquidada, y en los países más avanzados ha sido sustituida en los planes de estudio por la nueva ciencia lógica.

En los países de habla española sigue imperando Aristóteles; aunque no puede decirse que la nueva ciencia sea desconocida, es conocida en grado muy deficiente. Existe alguna traducción de Tarski y algún pequeño tratado escrito directamente en español, pero, en términos generales, puede afirmarse que los grandes de la lógica simbólica, los Boole, de Morgan, Peirce, Frege, Peano, Russell y Whitehead (por lo que hace a sus Principia Mathematica), Carnap, Tarski, etc., no han sido, en general, traducidos, o no lo han sido con la suficiente amplitud.

El libro que hoy ofrece la Compañía Editorial Continental, S. A., está, fundamentalmente, destinado al aprendizaje y, por lo mismo, se multiplica su utilidad en la tarea de llenar el vacío existente.

Es ya inaplazable la tarea de llevar a la escuela la lógica simbólica. La docencia científica necesita en nuestro medio ajustarse al ritmo de una actividad económica y social basada en las más modernas técnicas de computación electrónica o continuará siempre ese medio social necesitado de ayuda técnica. Junto a esa actividad de raigambre científica ultramoderna, se escucha aún en la vetusta cátedra de lógica la voz del maestro que demuestra "irrefutablemente" el error en que se encuentran la lógica de tendencia epistemológica, la lógica de tipo psicológico ... y la lógica simbólica.

El editor inglés sostiene que este libro contiene aportaciones originales a la ciencia de la lógica matemática. El autor, en efecto, hace una exposición sumamente original y, sobre todo, conducente al manejo efectivo de la disciplina, esto es, da preferencia a la técnica. Podría decirse que, en cierto modo, este libro es el reverso de la medalla de la obra de Susan K. Langer, que manifiestamente quiso escribir más sobre los fundamentos que sobre las "tretas" (como ella dice) de la materia; en cambio, el autor, me parece, insiste más en las "tretas" que en los fundamentos, lo que, naturalmente, hace más valioso este libro en nuestro medio que, por el momento, necesita, antes que otra cosa, el dominio de las técnicas de la lógica simbólica.

Tenemos, pues, la firme esperanza de que esta obra sea un eficaz instrumento para promover en los pueblos de habla española la expansión del conocimiento de una disciplina cuyo advenimiento a los primeros planos de la docencia es una necesidad de nuestro tiempo.

G. A. C.

ACLARACION IMPORTANTE

En esta traducción, cuando la letra "y" se utiliza como símbolo, suena "ye" y sólo cuando es una conjunción copulativa tiene sonido de i latina; a eso se debe que, en vez de expresiones como esta: x e y, se encuentren estas otras: x y y.

CONTENIDO

CAP. PÁG.

PREFACIO 7

PRESENTACIÓN DE LA EDICIÓN EN ESPAÑOL 11

PARTE I—PRINCIPIOS DE INFERENCIA Y DEFINICION

1. LOS CONECTIVOS ORACIONALES 25
   - 1.1 Negación y conjunción 25
   - 1.2 Disyunción 28
   - 1.3 Implicación: oraciones condicionales 29
   - 1.4 Equivalencia: oraciones bicondicionales 33
   - 1.5 Agrupamiento y paréntesis 33
   - 1.6 Tablas de verdad y tautologías 34
   - 1.7 Implicación y equivalencia tautológicas 39

2. TEORÍA ORACIONAL DE LA INFERENCIA 45
   - 2.1 Dos criterios principales de la inferencia e interpretaciones oracionales 45
   - 2.2 Las tres reglas oracionales de la deducción 51
   - 2.3 Algunas implicaciones tautológicas útiles 59
   - 2.4 Consistencia de las premisas y pruebas indirectas .. 64

3. REPRESENTACIÓN SIMBÓLICA DEL LENGUAJE COTIDIANO .. 73
   - 3.1 Gramática y lógica 73
   - 3.2 Términos 74
   - 3.3 Predicados 76
   - 3.4 Cuantificadores 78
   - 3.5 Variables ligadas y libres 83
   - 3.6 Un ejemplo final 87

CAP. PÁG.

4. TEORÍA GENERAL DE LA INFERENCIA ... 91
4.1 La inferencia, cuando implica sólo cuantificadores universales ... 91
4.2 Interpretaciones y validez ... 98
4.3 Inferencias restringidas con cuantificadores existenciales ... 116
4.4 Intercambio de cuantificadores ... 124
4.5 Inferencias generales ... 126
4.6 Resumen de las reglas de inferencia ... 137

5. OTRAS REGLAS DE INFERENCIA ... 141
5.1 Lógica de la identidad ... 141
5.2 Teoremas de lógica ... 149
5.3 Reglas deducidas de inferencia ... 155

6. APOSTILLA SOBRE EL USO Y LA MENCIÓN ... 163
6.1 Nombres y cosas nombradas ... 163
6.2 Problemas de las variables oracionales ... 165
6.3 Yuxtaposición de nombres ... 167

7. TRANSICIÓN DE LAS DEMOSTRACIONES FORMALES A LAS INFORMALES ... 171
7.1 Consideraciones generales ... 171
7.2 Axiomas básicos sobre números ... 172
7.3 Ejemplos comparativos de deducciones formales y de demostraciones informales ... 175
7.4 Ejemplos de demostraciones informales falaces ... 182
7.5 Otros ejemplos de demostraciones informales ... 187

8. TEORÍA DE LA DEFINICIÓN ... 197
8.1 Ideas tradicionales ... 197
8.2 Criterios para definiciones adecuadas ... 198
8.3 Reglas de las definiciones adecuadas ... 202
8.4 Definiciones que son identidades ... 208
8.5 El problema de la división entre cero ... 210
8.6 Definiciones condicionales ... 212
8.7 Cinco métodos de la división por cero ... 213
8.8 Principio de Padoa e independencia de los símbolos primitivos ... 217

CAP. PÁG.

PARTE II—TEORIA ELEMENTAL INTUITIVA DE LOS CONJUNTOS

9. CONJUNTOS 225
9.1 Introducción 225
9.2 Membrecía 225
9.3 Inclusión 229
9.4 El conjunto vacío 232
9.5 Operaciones con conjuntos 233
9.6 Dominios de individuos 236
9.7 Transcripción del lenguaje cotidiano 238
9.8 Diagramas de Venn 245
9.9 Principios elementales de las operaciones con conjuntos 252

10. RELACIONES 259
10.1 Pares ordenados 259
10.2 Definición de las relaciones 261
10.3 Propiedades de las relaciones binarias 264
10.4 Relaciones de equivalencia 270
10.5 Relaciones de ordenación 272
10.6 Operaciones con relaciones 278

11. FUNCIONES 283
11.1 Definición 283
11.2 Operaciones con las funciones 289
11.3 Notación lambda de Church 297

12. FUNDAMENTACIÓN SINFOREMÁTICA DEL MÉTODO AXIOMÁTICO 303
12.1 Introducción 303
12.2 Predicados sinforemáticos y axiomatizaciones de teorías 307
12.3 Isomorfismo de modelos para una teoría 320
12.4 Ejemplo: Probabilidad 335
12.5 Ejemplo: Mecánica 356

INDICE ANALÍTICO 373

INTRODUCCION

Nuestro uso cotidiano del lenguaje es vago y nuestro nivel diario de pensamiento es, a menudo, confuso. Uno de los principales propósitos de este libro es el de iniciar al lector en una forma de pensamiento que propicia el esmero y la precisión. Hay muchos modos de aprender cómo usar con precisión el lenguaje y las ideas. Nuestro método se basará en el estudio de la lógica. En los tiempos que corren, la lógica se ha convertido en una disciplina amplia y profunda. Para comenzar, nos concentraremos en la parte que trata de la teoría del razonamiento correcto, también llamada teoría de la inferencia lógica, teoría de la demostración, o teoría de la deducción. Los principios de la inferencia lógica son universalmente aplicados en todas las ramas del conocimiento sistemático. Se dice a menudo que la prueba más importante de cualquier teoría científica, radica en su utilidad y exactitud en la descripción de fenómenos, antes de que estos ocurran. Una predicción de esta naturaleza debe involucrar la aplicación de los principios de la inferencia lógica. Por ejemplo, si sabemos cuáles son las fuerzas que actúan sobre un cuerpo y sabemos dónde está el cuerpo en un momento dado y cuál es su velocidad, podemos usar la teoría de la mecánica, juntamente con las reglas de la inferencia lógica y ciertos teoremas de matemáticas, para predecir dónde estará el cuerpo cierto tiempo después.

Durante dos mil años los matemáticos han estado haciendo inferencias correctas, de índole sistemática e intrincada, y los lógicos y filósofos han estado analizando la naturaleza de los argumentos válidos. Es, en consecuencia, sorprendente que sólo en las tres o cuatro últimas décadas haya sido desarrollada una teoría formal de la inferencia plenamente satisfactoria. En el largo periodo que se extiende desde Aristóteles, en el siglo IV A. C., hasta Leibniz en el siglo XVII, los lógicos de la antigüedad, los medievales y los postmedievales descubrieron mucho de importancia y significación sobre la lógica, pero el más importante defecto de esta tradición clásica era su fracaso en relacionar la lógica como teoría de la inferencia a

la clase de razonamientos deductivos que se usan continuamente en matemáticas.

Leibniz tenía cierta visión sobre la necesidad de hacer esta conexión, pero hasta fines del siglo XIX y comienzos del presente se establecieron relaciones sistemáticas entre la lógica y las matemáticas, principalmente mediante las obras de Frege, Peano y Russell. A pesar del alcance y magnitud de sus investigaciones, sólo en años recientes ha sido formulada una teoría completamente explícita de la inferencia, adecuada para manejar todos los casos paradigmáticos de razonamiento deductivo, en matemáticas y en las ciencias empíricas. Es grande el número de personas que han contribuido a estos desarrollos recientes, pero quizás los más prominentes han sido Kurt Gödel, David Hilbert y Alfredo Tarski.

No obstante, sería un error pensar que la teoría de la inferencia desarrollada en la primera parte de este libro tiene entidad solamente en lo que respecta a contextos científicos. La teoría se aplica tanto a actuaciones en tribunales como a análisis filosóficos de las verdades eternas. En realidad, no es demasiado pretender que la teoría de la inferencia atañe a toda seria deliberación humana.

Un razonamiento correcto, ya sea en matemáticas, física, o en la conversación casual, es válido en virtud de su forma lógica. En virtud de que la mayor parte de los argumentos se expresan en lenguaje usual, con adición de unos cuantos símbolos pertinentes a la disciplina de que se trate, no resulta transparente la forma lógica del argumento. Afortunadamente, es posible poner de manifiesto esta estructura lógica, aislando un pequeño número de palabras y frases clave, como "y", "no", "todo" y "algunos". A fin de hacer precisiones sobre estas expresiones centrales y de establecer reglas explícitas de inferencia, basadas en la incidencia de las mismas, uno de nuestros primeros pasos será el de asignarles símbolos lógicos. Con ayuda de estos símbolos es relativamente fácil enunciar y aplicar reglas de inferencia válida, tarea que ocupa los primeros siete capítulos.

Para traer precisión lógica a nuestro análisis de las ideas, no es, por lo general, suficiente con construir inferencias válidas; es también esencial tener cierto dominio de los métodos para definir en forma exacta un concepto, en función de otros conceptos. En cualquier rama dada de la ciencia o de las matemáticas, uno de los métodos más poderosos para eliminar la vaguedad conceptual es el de aislar un pequeño número de conceptos básicos para el sujeto de que se trate, y a continuación definir los demás conceptos de la disciplina, en función del conjunto básico. El propósito del Cap. 8 es el de establecer reglas exactas para dar esas definiciones. Se demostrará que tanto las definiciones correctas como las inferencias

correctas dependen primordialmente de cuestiones de forma lógica. Sin embargo, se suscitan ciertas sutiles cuestiones de existencia en la teoría de la definición, que carecen de contrapartida en la teoría de la inferencia.

Los primeros ocho capítulos constituyen la Parte I, que está dedicada a los principios generales de la inferencia y la definición. La Parte II, los cuatro últimos capítulos, trata de la teoría elemental de los conjuntos. A causa de que los varios respectos en que esta teoría está íntimamente ligada con la lógica, no resultarán familiares para muchos lectores, no estará de más una explicación sobre la inclusión de este material.

La teoría de los conjuntos, o la teoría general de las clases, como se llama algunas veces, es la disciplina básica de las matemáticas, pues con unas cuantas raras excepciones, las entidades que se estudian y analizan en matemáticas pueden considerarse como ciertos conjuntos o clases particulares de objetos. Como veremos, los objetos estudiados en una rama de las matemáticas puras, como la teoría de los conjuntos, o en una rama de la física matemática, como la teoría de la mecánica, pueden caracterizarse como ciertos conjuntos. Por esta razón cualquier parte de las matemáticas puede considerarse como una rama especial de la teoría de los conjuntos. Sin embargo, como esta interpretación identificaría la teoría de los conjuntos con la totalidad de las matemáticas, se acostumbra reservar la expresión "teoría de los conjuntos" para la teoría general de las clases o conjuntos y para ciertos tópicos, como la construcción de los enteros y números reales como conjuntos que, históricamente, están estrechamente relacionados con investigaciones sobre los fundamentos de las matemáticas.

El primer capítulo de la Parte II contiene una exposición intuitiva de las más importantes relaciones existentes entre conjuntos arbitrarios. Hay, por ejemplo, operaciones simples sobre conjuntos que corresponden a las operaciones aritméticas de adición, multiplicación y sustracción. El siguiente capítulo (Cap. 10), trata de la teoría de las relaciones, que se incorpora a la teoría de los conjuntos por medio del concepto de un par ordenado de objetos. Se hace hincapié en las relaciones de ordenación, a causa de su importancia en muchas ramas de las matemáticas y la ciencia. El Cap. 11 trata de las funciones que, desde el punto de vista de la teoría de los conjuntos, son únicamente relaciones poseedoras de una propiedad especial.

Mientras que los primeros tres capítulos de la Parte II están dedicados a la teoría de los conjuntos, el capítulo final (Cap. 12), se ocupa de la relación entre la teoría de los conjuntos y ciertas cuestiones metodológicas o fundamentales de matemáticas y de fi-

losofía. El tema central de este capítulo es el de indicar cómo puede fundamentarse axiomáticamente en la teoría de los conjuntos cualquier rama de matemáticas o cualquier teoría científica. El punto de vista que se expone en detalle en el Cap. 12 es que el mejor modo de fundamentar axiomáticamente una teoría es definiendo un predicado adecuado dentro de la teoría de los conjuntos.

Desde principios de este siglo, los filósofos han escrito mucho sobre la estructura de las teorías científicas, pero han dicho lamentablemente poco sobre la estructura detallada de cada teoría en particular. La fundamentación axiomática de una teoría dentro de la teoría de los conjuntos es un importante paso inicial en lo que respecta a darle una estructura exacta y explícita. Por ejemplo, ¿cuándo son isomórficos, esto es, cuándo tienen exactamente la misma estructura dos modelos de una teoría? En realidad los problemas filosóficos bien conocidos, como la reducción de una rama de las ciencias empíricas a otra, pueden ganar precisión en función de esos conceptos de la teoría de los conjuntos, como el de isomorfismo. En el capítulo doce se dan aplicaciones de estas ideas a ejemplos sustantivos de las matemáticas puras y las ciencias empíricas.

La finalidad de las partes primera y segunda es la de presentar la lógica como parte de las matemáticas y la ciencia y la de mostrar, por medio de numerosos y pormenorizados ejemplos, la importancia de la lógica, aun para ciencias empíricas, como la psicología. Por esta razón puede decirse que el centro de gravedad de este libro se ha colocado en el uso y la aplicación sistemáticos de la lógica, más bien que en el desarrollo de la misma como una disciplina autónoma.

Finalmente, debe observarse que no se intenta presentar en estas páginas ninguna definición precisa de lógica. En un sentido restringido, lógica es la teoría de los argumentos válidos o la teoría de la inferencia deductiva. Un sentido ligeramente más amplio incluye la teoría de la definición y otro aún más amplio incluye la teoría general de los conjuntos. Además, la teoría de la definición, juntamente con la de los conjuntos, suministra una fundamentación exacta para el método axiomático, cuyo estudio consideran informalmente la mayor parte de los matemáticos como parte de la lógica.

PARTE I

PRINCIPIOS DE INFERENCIA Y DEFINICION

CAPITULO 1

LOS CONECTIVOS* ORACIONALES

Para comenzar, deseamos desarrollar un vocabulario que sea preciso y al mismo tiempo adecuado para el análisis de los problemas y conceptos del conocimiento sistemático. Debemos usar un lenguaje vago para crear un lenguaje preciso. Esto no es tan tonto como parece. Las reglas de ajedrez, por ejemplo, son mucho más precisas que las de la gramática y, no obstante, usamos oraciones regidas por reglas gramaticales imprecisas, para enunciar la precisas reglas del ajedrez. De hecho, nuestro primer paso será un tanto semejante al de establecer las reglas de un juego. Deseamos establecer reglas estrictas sobre el uso de ciertas palabras clave: "no", "y", "o", "si... entonces ....", "si, y sólo si", que se llaman *conectivos oracionales*. Las reglas para su uso, sin embargo, no tendrán el carácter de reglas de un juego arbitrario. Serán diseñados para hacer explícito el uso sistemático predominante de estas palabras; este uso sistemático ha sido puesto de manifiesto por la reflexión sobre las formas en que se usan en los contextos cotidianos y ordinarios. Sin embargo, no dudaremos en desviarnos del uso ordinario, cada vez que existan razones de peso para hacerlo.

§1.1. Negación y conjunción. Negamos la verdad de una oración afirmando su negación. Por ejemplo, si pensamos que la oración "El azúcar produce caries dental" es falsa, afirmamos la oración "El azúcar no produce caries dental". En este ejemplo se ilustra el método usual de aseverar la negación de una oración simple: fijamos la palabra "no" al verbo principal de la oración. En cambio, la aserción de la negación de una oración compuesta es más compli-

* Al avanzar en la lectura del texto se hará evidente la necesidad de aceptar el neologismo, consistente en el uso como sustantivo de una forma fundamentalmente adjetival. Existe, sin embargo, el antecedente kantiano de "imperativo", por mandato u ordenamiento. (N. del T.)

cada. Por ejemplo, negamos la oración "El azúcar produce caries dental y el whiskey úlceras", afirmando, "No es cierto que el azúcar produzca caries dental y el whiskey úlceras". A pesar de la aparente divergencia entre estos dos ejemplos, es conveniente adoptar en lógica un solo signo para simbolizar la negación de una oración. Usaremos el prefijo '—', que se coloca antes de la oración, considerada como un todo. Así pues, la negación del primer ejemplo se escribe:

—(El azúcar produce caries dental).

El segundo ejemplo ilustra la forma en que podemos siempre interpretar el símbolo '−'; podemos usar siempre la expresión 'no es cierto que', o 'no es verdad que'.

La principal razón para adoptar el símbolo único '−' para la negación, sin tener en cuenta que la sentencia negada sea simple o compuesta, es que el significado del signo es el mismo en ambos casos. *La negación de una oración verdadera es falsa, y la negación de una oración falsa es verdadera.*

Usamos la palabra 'y' para conjuntar* dos oraciones, a fin de hacer una sola oración que llamamos la *conjunción* de las dos oraciones. Por ejemplo, la oración 'María ama a Juan y Juan ama a María' es la conjunción de la oración 'María ama a Juan', con la oración 'Juan ama a María'. Para simbolizar la conjunción, usaremos el signo '&'. Así pues, la conjunción de dos sentencias cualesquiera, **P** y **Q**, se escribe

P & Q

La regla que rige el uso del signo '&' está completamente de acuerdo con el uso ordinario. *La conjunción de dos oraciones es verdadera si, y sólo si ambas oraciones son verdaderas.* Hay que hacer notar que en lógica podemos combinar dos oraciones cualesquiera para formar una conjunción. No se requiere que las dos oraciones sean afines en su contenido o materia de que traten. Cualesquiera combinaciones, por absurdas que sean, son permitidas. Por supuesto, habitualmente no nos interesan oraciones como 'Juan ama a María y 4 es divisible por 2'. Aunque pudiera parecer deseable contar con una regla adicional enunciativa de que sólo podemos conjuntar ora-

* De nuevo se impone el neologismo "conjuntar", en vez del aceptado "conjugar" que tiene tradicionalmente un significado gramatical que lo incapacita para nuestro propósito. Por lo demás, conjuntar tendría como simple correspondiente al vocablo castizo "juntar"; conjugar, en cambio, es un cultismo que carece de simple. Que conjugar sea un cultismo es obvio por su etimología (conjugo, avi, atum, are), y que juntar sea castizo lo indica el mismo criterio (jungo, junxi, junctum) que insinúa un verbo del latín vulgar, "junctare". Tiene, pues, más derechos, por su prosapia castiza, "conjuntar" que conjugar. (N. del T.)

ciones de contenido común, lo indeseable de esa regla resulta evidente cuando reflexionamos sobre la vaguedad de la idea de "contenido común".

Se usan varias palabras como sinónimos aproximados de 'no' y de 'y', en el lenguaje ordinario. Por ejemplo, la palabra 'nunca' en la oración:

Nunca accederé a tus exigencias

tiene casi el mismo significado que 'no' en

No accederé a tus exigencias.

No obstante, es verdad que 'nunca' es portadora de un significado de negación continuada de que carece 'no'.

La palabra 'pero' tiene más o menos el mismo sentido de 'y' y la simbolizamos con '&', aunque en muchos casos del uso común hay diferencias de significado. Por ejemplos, si una joven dice a un joven:

Te amo y amo casi lo mismo a tu hermano.

el joven reaccionaría en forma diferente que si le dijera:

Te amo, pero amo casi lo mismo a tu hermano.

En vista de esas diferencias de significado, parece muy natural la idea de introducir símbolos diferentes para los conectivos oracionales 'nunca' y 'pero'. Existe, sin embargo, un sólido argumento en contra. Las reglas convenidas sobre el uso o la aplicación de la negación y la conjunción hacen que estos dos conectivos oracionales resulten *aleticofuncionales** o *funcionales con respecto de la verdad o de lo verdadero*, esto es, la verdad o la falsedad de la negación de una oración **P**, o la verdad o la falsedad de la conjunción de dos oraciones **P** y **Q** es precisamente una función de la verdad o falsedad de **P** en el caso de negación y de **P** y **Q** en el caso de conjunción. Está claro que queda descartado un análisis aleticofuncional de 'pero', distinto del dado para 'y'; aventurarse, por otra parte, en un análisis no aleticofuncional, lleva a conclusiones vagas y oscuras. Cualquier duda sobre este particular es rápidamente disipada si se hace el intento de establecer una regla precisa sobre el uso de 'pero' que difiera de la que ya se ha dado para 'y'.

Por supuesto, que la rica y matizada naturaleza de los idiomas garantiza que en muchos contextos pueden ser usados los conectivos en formas no aleticofuncionales, con delicados matices de sig-

* Hemos traducido *truth-functional* como "aleticofuncional y orgánico"; desde luego, esta segunda expresión es preferible, por no ser híbrida como la primera. (N. del T.)

nificación. La pérdida de sutileza en el matiz, parece acompañar inevitablemente a la formulación de un preciso análisis simbólico de las oraciones. Pero este proceso de abstracción deformadora no es privativo de la lógica; es una característica de la ciencia en general. Pocos poetas estarían interesados en un análisis aleticofuncional del lenguaje, y ningún naturalista tomaría en cuenta los conceptos del físico, de posición, velocidad, aceleración, masa y fuerza, como adecuados para describir el vuelo de un águila. Los conceptos de lógica desarrollados en este libro son útiles para descubrir y comunicar el conocimiento sistemático, pero su pertinencia con respecto de otras funciones del lenguaje y del pensamiento es menos directa.

§1.2. Disyunción. Usamos la palabra 'o' para efectuar la *disyunción* de dos oraciones. En el lenguaje usual se usa la palabra o en dos sentidos diferentes. En el llamado sentido *no exclusivo* la disyunción de dos oraciones es verdadera si cuando menos una de las dos es verdadera. En los contratos, este sentido se expresa a menudo por medio del barbarismo 'y/o', ilustrado en el siguiente ejemplo:

> Antes de que se haga el trabajo mencionado o se suministren los materiales que se indican, el arrendatario y cualquier contratista o persona comisionada para hacer ese trabajo y/o suministrar los materiales indicados, otorgará la caución o cauciones que el arrendador estime pertinente y no lesiva......

Hacemos notar que en el ejemplo anterior no hay disyunciones de oraciones, sino de cláusulas o expresiones que no son oraciones. Encontraremos, sin embargo, que es más conveniente tratar esos ejemplos como disyunciones de oraciones; este punto de vista refleja otra divergencia entre la lógica y el lenguaje usual.

La palabra latina 'vel' tiene aproximadamente el significado de 'o' en el sentido no exclusivo y, consecuentemente, usamos el signo 'V' para la disyunción de dos oraciones en este sentido. Así pues, la disyunción de dos oraciones cualesquiera **P** y **Q** se escribe

$$P \vee Q.$$

Restringiremos el uso de la palabra 'disyunción' al sentido no exclusivo y nuestra regla sobre su uso es: *La disyunción de dos oraciones es verdadera si, y sólo si cuando menos una de las oraciones es verdadera.*

Cuando se usa 'o' en el sentido *exclusivo* para combinar dos oraciones, se enuncia que una de las oraciones es verdadera y la otra es falsa. Esta significación se hace a menudo más explícita añadiendo 'ambos' o 'los dos', en el caso gramatical que corresponda, y

en la oración elíptica y subordinada 'pero no ambos' o 'pero no los dos'. Por ejemplo, un padre dice a su hijo: 'Puedes ir al cine o al circo este sábado, pero no a los dos'. No introduciremos un signo especial para 'o' en el sentido exclusivo, pues resulta que en las discusiones científicas siempre es posible arreglárselas con 'o' en el sentido no exclusivo (que también se llama el sentido *inclusivo*).

§1.3. Implicación:* oraciones condicionales. Usamos las palabras 'si . . . , entonces . . . ' para obtener de dos oraciones una *oración condicional.* Una oración condicional se llama también una *implicación.* En el significado que tienen las palabras en el lenguaje cotidiano, es difícil precisar en qué circunstancias la mayor parte de las personas acepta como verdadera una oración condicional. Considérese un ejemplo similar a otro que ya se ha usado:

(1) Si María ama a Juan, entonces Juan ama a María.

Si la oración 'María ama a Juan' es verdadera y la oración 'Juan ama a María' es falsa, entonces todos convendrían en que (1) es falsa. Además, si la oración 'María ama a Juan' es verdadera y la oración 'Juan ama a María' es también verdadera, entonces casi todos convendrían en que (1) es verdadera. Las dos posibilidades de verdad y falsedad que acabamos de enunciar son las únicas que se suscitan muy a menudo en el uso ordinario del lenguaje. Hay, no obstante, dos posibilidades más, y si interrogamos sobre ellas al proverbial hombre común y corriente, es imposible conjeturar su respuesta. Estos dos casos adicionales son como sigue. Supóngase que la oración 'María ama a Juan' es falsa; en ese caso, ¿qué tenemos que decir sobre la verdad de (1): primero, cuando la verdadera oración 'Juan ama a María' es también falsa, y, segundo, cuando la oración 'Juan ama a María' es verdadera? En matemáticas y en lógica esta cuestión se contesta en la forma siguiente: la oración (1) es verdadera si la oración 'María ama a Juan' es falsa, independientemente de la verdad o falsedad de la oración 'Juan ama a María'.

Para enunciar nuestra regla sobre el significado de 'si . . . entonces . . . ' es conveniente usar una terminología especial, llamando a la oración que sigue inmediatamente después de 'si' el *antecedente* o *hipótesis* de la oración condicional y *consecuente* o *conclusión* a la oración que sigue inmediatamente después de 'entonces'. Así pues, 'María ama a Juan' es el antecedente de (1) y 'Juan ama a María' su consecuente. La regla de significado es, pues: *Una oración condicional es falsa si el antecedente es verdadero y el consecuente es falso; en cualquier otro caso es verdadera.*

* Se trata de lo que la lógica aristotélica ha llamado "inferencias inmediatas", aunque no hay una correspondencia exacta. (N. del T.)

En dos respectos pueden hacerse objeciones intuitivas a esta regla. En primer lugar, puede sostenerse que la implicación no es un conectivo aleticofuncional, y que debe existir cierta conexión definida entre el antecedente y el consecuente de una oración condicional. De acuerdo con la regla de significado que se acaba de enunciar, la oración:

(2) Si la poesía es para los jóvenes, entonces $3 + 8 = 11$

es verdadera, puesto que el consecuente es verdadero. Y sin embargo, muchas personas estarían dispuestas a desechar esa oración como disparatada; alegarían que la verdad del consecuente de ninguna manera *depende* de la verdad del antecedente y que, en consecuencia, la oración (2) no constituye una implicación válida. Sin embargo, la adhesión del lógico a los conectivos aleticofuncionales no deja de tener sus razones. ¿Cómo va a caracterizarse una noción tan oscura como la de dependencia? Este es el mismo problema que se encontró al estudiar las conjunciones. Si acaso se piensa que se está evadiendo un problema importante, quizás decisivo, con la excusa de que es difícil, en el capítulo siguiente se darán pruebas de que los conectivos aleticofuncionales son muy adecuados tanto para la teoría como para la práctica de la inferencia lógica.

Aun cuando se acepten los postulados aleticofuncionales, una segunda objeción a la regla sobre el significado de la implicación es la de que se ha establecido un principio falso al llamar verdadera a una implicación cuando su antecedente es falso. Pero los ejemplos particulares arguyen enérgicamente en favor de esta regla. Para el caso en que el antecedente es falso, considérese:

(3) Si hay aproximadamente cien millones de esposos en los Estados Unidos, entonces hay aproximadamente cien millones de esposas en los Estados Unidos.

Es difícil imaginar que alguien niegue la verdad de (3). Para el caso en que el consecuente es verdadero, considérese la siguiente modificación de (3):

(4) Si hay aproximadamente cien mllones de esposos en los Estados Unidos, entonces el número de esposos en este país es mayor que el número de esposos en Francia.

Si se admite que (3) y (4) son verdaderas, entonces ha quedado establecida la regla *aleticofuncional* de las oraciones condicionales con antecedente falso.

Puede objetarse que escogiendo ejemplos ligeramente diferentes puede probarse que cualquier implicación es falsa cuando su antece-

dente es falso. Por ejemplo, supóngase que (3) fuera reemplazada por:

(5) Si hay aproximadamente cien millones de esposos en los Estados Unidos, entonces hay exactamente una esposa en los Estados Unidos.

Entonces, dentro de nuestro esquema aleticofuncional podría sostenerse que es falsa una implicación con falso antecedente y falso consecuente, puesto que podría argüirse plausiblemente que (5) es falsa en su acepción ordinaria. Sin embargo, existen buenos motivos para escoger a (3) de preferencia a (5), pues (3) tiene la propiedad de que su consecuente se sigue de su antecedente, de conformidad con algunos principios generales, aritméticos y conyugales. Por lo que respecta a (5), no parece posible establecer ese hilo intuitivo de razonamiento; de hecho, es precisamente la ausencia de esa conexión lo que nos hace declararla falsa. Aunque ya hemos admitido que la noción de conexión o dependencia a que aquí se recurre es demasiado vaga para constituir un concepto formal de la lógica, al escoger ejemplos que nos impongan la evidencia —dentro de nuestro esquema aleticofuncional— de la verdad de implicaciones con antecedentes falsos, es razonable escoger un ejemplo como (3), para el cual nuestra intuición de la dependencia es fuerte, y no un ejemplo como el (5), para el cual es débil. La exigencia aletico funcional de que se consideren verdaderas oraciones como (5) no tiene efectos indeseables, pues las oraciones condicionales cuyos antecedentes y consecuentes no están relacionados, y cuyos antecedentes son falsos, no desempeñan un papel importante en la argumentación sistemática.

En cuanto a notación, la oración condicional formada por dos oraciones **P** y **Q** se escribe:

$$\mathbf{P} \rightarrow \mathbf{Q}.$$

El signo '→' se llama a menudo signo de *implicación*. Algunas otras expresiones idiomáticas españolas tienen aproximadamente el mismo significado sistemático de 'si . . . , entonces . . .'. Representamos también con el símbolo **P** → **Q**, las expresiones

P solamente si **Q**
Q si **P**
Q siempre que **P**
P es condición suficiente de **Q**
Q es condición necesaria de **P**

De estas cuatro expresiones, es muy notorio el uso variante de 'solamente si'. Es un 'error' muy divulgado el uso de 'solamente si' en el sentido de 'si'. Por ejemplo, la oración:

(6) Juan lleva a pasear a María solamente si Isabel está disgustada con él

no tendría ordinariamente el mismo significado que

Si Juan lleva a pasear a María, entonces Isabel está disgustada con él,

y sería más precioso (pero todavía no del todo correcto idiomáticamente), interpretar (6) en el sentido de

Si Isabel está disgustada con él, entonces Juan lleva a pasear a María.

La predominancia de oraciones como (6) hace difícil para muchas personas que por primera vez estudian lógica o matemáticas aceptar la estipulación* de que

(7) **P** solamente si **Q**

significa lo mismo que

(8) Si **P**, entonces **Q**.

Y no obstante, es cierto que en el razonamiento científico (7) y (8) son idiomáticamente equivalentes y como tales se tratarán en todo lo que resta de este libro.

Por lo que respecta a las dos últimas expresiones idiomáticas, convienen hacer notar que son ampliamente usadas en matemáticas. Así, la oración 'si un triángulo es equilátero, entonces es isósceles', puede ser nuevamente formulada como sigue:

Para que un triángulo sea isósceles, es suficiente con que sea equilátero,

o

Es necesario que un triángulo equilátero sea isósceles.

Nótese que proceden algunos cambios gramaticales en las oraciones componentes **P** y **Q**, cuando se pasa de

Si **P**, entonces **Q**

a

(9) **P** es una condición suficiente de **Q**,

* Se usa la palabra "estipulación" en un sentido distinto al de "convención"; se entiende que la estipulación no es arbitraria como la convención o, cuando menos, no del todo. (N. del T.)

de manera que (9) no es una formulación exacta; pero estos cambios son generalmente obvios y no es necesario continuar aquí su examen.

§1.4. Equivalencia: oraciones bicondicionales. Usamos la expresión 'si, y sólo si' para obtener de dos oraciones una *oración bicondicional*. Una oración bicondicional se llama también una *equivalencia* y las dos oraciones conectadas por medio de 'si, y sólo si' se llaman, respectivamente, el *primero* y el *segundo miembro de la equivalencia*. La bicondicional

(1) **P**, si, y sólo si, **Q**

Tiene el mismo significado que la oración

(2) **P** si **Q**, y **P**, si solamente, **Q**

y (2) es equivalente a

(3) Si **P**, entonces **Q**, y si **Q**, entonces **P**.

Nuestras reglas de significado de la conjunción y la implicación nos dicen que (3) es verdadera precisamente cuando **P** y **Q** son ambas verdaderas o ambas falsas. De aquí la regla: *Una oración bicondicional es verdadera, si, y sólo si, sus dos miembros son ambos verdaderos o ambos falsos.* Por lo que hace a la notación, se escribe

$$P \leftrightarrow Q$$

para la bicondicional formada por las oraciones **P** y **Q**.

De acuerdo con las observaciones que hicimos al final del último apartado, debe hacerse notar que (1) es equivalente a

Q es una condición necesaria y suficiente para **P**.

§1.5. Agrupamiento y paréntesis. En el lenguaje ordinario el agrupamiento adecuado de oraciones que se combinan en una oración compuesta se indica por una variedad de medios o dispositivos lingüísticos. Cuando se simbolizan estas oraciones en lógica, los mencionados dispositivos pueden consistir en un adecuado uso de paréntesis.

Por ejemplo, la oración:

Si Showboat gana la carrera, entonces lo indicarán Shotless y Ursula, se simboliza por

(1) $S \rightarrow (H \mathbin{\&} U)$,

donde **S** es 'Showboat gana la carrera', **H** es 'Shotless lo indicará' y **U** es 'Ursula lo indicará'. Leemos (1)

Si **S**, entonces **H** y **U**.

Leemos, por otra parte,

(2) $(\mathbf{S} \rightarrow \mathbf{H})\ \&\ \mathbf{U}$,

como

Ambas cosas: si **S** entonces **H**, y **U**.

De este modo debe resultar claro que (1) más bien que (2) es la simbolización correcta de la oración original. Los paréntesis se usan de un modo natural, que resulta familiar del álgebra elemental, para indicar cuál conectivo es el dominante.

Si se adopta una convención natural con respecto de la dominancia relativa de los diversos conectivos, se obtendrá una considerable reducción de los paréntesis usados en la práctica. La convención es '↔' y '→' dominan a '&' y 'V'. De este modo (1) puede escribirse

(3) $\mathbf{S} \rightarrow \mathbf{H}\ \&\ \mathbf{U}$

y

$$\mathbf{P} \leftrightarrow \mathbf{Q}\ \&\ \mathbf{R}$$

significa

$$\mathbf{P} \leftrightarrow (\mathbf{Q}\ \&\ \mathbf{R}).$$

En cambio, si rige esta convención, no está claro lo que significa

$$\mathbf{P}\ \&\ \mathbf{Q} \vee \mathbf{R}$$

y lo mismo puede decirse de

$$\mathbf{P} \leftrightarrow \mathbf{Q} \rightarrow \mathbf{R}.$$

§1.6. Tablas de verdad y tautologías. Nuestras reglas aletico-funcionales sobre el uso de la negación, la conjunción, la disyunción, la implicación y la equivalencia, pueden resumirse en forma tabular. Estas *tablas básicas de verdad* nos dicen a primera vista en qué circunstancias es verdadera la negación de una oración, si conoce mos la verdad o falsedad de ésta, y de modo semejante para la conjunción, la disyunción o la implicación de dos oraciones.

Negación*

P	-P
T	F
F	T

Conjunción

P	Q	P & Q
T	T	T
T	F	F
F	T	F
F	F	F

Disyunción

P	Q	P ∨ Q
T	T	T
T	F	T
F	T	T
F	F	F

* En este caso, y en todos los demás, hemos respetado el simbolismo del autor. Dentro de esa norma, usamos *F* para falso y *T* para verdadero. Entre los símbolos *V* y *T* para

Implicación		
P	Q	P → Q
T	T	T
T	F	F
F	T	T
F	F	T

Equivalencia		
P	Q	P ↔ Q
T	T	T
T	F	F
F	T	F
F	F	T

Podemos planear del siguiente modo el uso de las tablas de verdad. Si **N** es la oración verdadera 'Newton nació en 1642' y **G** es la oración falsa 'Galileo murió en 1640', podemos entonces calcular la verdad o falsedad de una oración compuesta tan complicada como

(1) $((\mathbf{N} \vee \mathbf{G}) \ \& \ —\mathbf{N}) \rightarrow (\mathbf{G} \rightarrow \mathbf{N})$.

Puesto que **N** es verdadera, vemos por la tabla de la disyunción que **N** V **G** es verdadera; de la tabla de la negación que —**N** es falsa y de aquí, por la tabla de la conjunción, que el antecedente de (1) es falso. Finalmente, conforme a la tabla de la implicación concluimos que toda la oración es verdadera. En el siguiente análisis diagramático, que se explica por sí mismo, se ilustra una aplicación más explícita de las tablas de verdad, en forma análoga al uso de una tabla de multiplicar.

$((\mathbf{N} \vee \mathbf{G}) \ \& \ -\mathbf{N}) \rightarrow (\mathbf{G} \rightarrow \mathbf{N})$

(Nótese que en este diagrama el análisis procede de adentro hacia afuera. La última curva conecta los dos miembros del conectivo mayor).

Llamamos *atómica* a una oración si no contiene conectivos oracionales. Por ejemplo, la oración.

El Sr. Knightley amaba a Emma

es atómica, mientras que las oraciones

Emma no amaba a Frank Churchill

verdadero, no hay más diferencia sino que el primero es nemónico y el segundo arbitrario, pero nada obliga a usar símbolos nemónicos, por una parte, y, por otra, encerrarse dentro de un simbolismo particular para cada idioma es renunciar al viejo y justificado ideal de universalidad simbólica de las ciencias exactas. (N. del T.)

y

La Sra. Elton era una cursi y la Srita. Bates un fastidio;

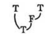
no son atómicas, pues la primera contiene una negación y la segunda una conjunción.

Usamos ahora el concepto de oración atómica para definir lo que probablemente es la más importante noción de este capítulo. La idea intuitiva es *que una oración compuesta es una tautología si es verdadera independientemente de los valores de verdad de sus oraciones atómicas componentes*. Por ejemplo, para cualquier oración atómica **P**,

P ∨ -P

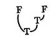
es una tautología. Si **P** es verdadera, tenemos

P ∨ -P



Si **P** es falsa, tenemos:

P ∨ -P



Así pues, sea **P** verdadera o falsa, **P V —P** es verdadera y, en consecuencia, una tautología. Las *tablas de verdad derivadas* son más cómodas y compactas que el análisis diagramático mostrado anteriormente, cuando deseamos saber si una oración es una tautología

P	-P	P ∨ -P
T	F	T
F	T	T

La segunda columna se deduce de la primera usando la tabla de negación y la tercera de las primeras dos usando la de disyunción. Puesto que ambas líneas de la columna final tienen la anotación 'T' toda la oración es una tautología. La idea de la tabla de verdad derivada es que una oración es una tautología si es verdad para todas

las combinaciones* posibles de valores de verdad de sus oraciones atómicas componentes. El número de esas combinaciones depende del número de oraciones atómicas componentes. Por ejemplo, si hay tres oraciones atómicas distintas, hay ocho combinaciones de valores de verdad posibles, pues cada oración atómica tienen exactamente dos valores de verdad posibles: verdad y falsedad. En general, si hay n oraciones atómicas componentes, hay 2^n combinaciones* de valores de verdad posibles, lo que significa que la tabla de verdad derivada de una oración compuesta de n oraciones atómicas distintas tiene 2^n líneas. Por ejemplo, para demostrar que **P V Q → P** no es una tautología cuando **P** y **Q** son oraciones atómicas distintas, necesitamos $2^2 = 4$ líneas, como en la siguiente tabla de verdad:

P	Q	P V Q	P V Q → P
T	T	T	T
T.	F	T	T
F	T	T	F
F	F	F	T

En esta tabla se obtiene la tercera columna deduciéndola de las dos primeras mediante el uso de la tabla de disyunción y la columna final se deduce de las columnas tercera y primera, aplicando la tabla de implicación. Puesto que en la tercera línea de la cuarta columna existe la anotación 'F' de falsa, se concluye que **P V Q → P** no es una tautología, pues esta tercera línea demuestra que si **P** es falsa y **Q** es verdadera, entonces **P V Q → P** es falsa. *Debe hacerse hincapié en que una oración es una tautología si, y solamente si, todas las anotaciones de la columna final son iguales a 'T' (verdadera).* Basta con que aparezca la letra 'F' en una sola línea de la columna final para garantizar que la oración que se analiza no es una tautología.

Cuando postulamos como nuestra definición formal:

> *Una oración es una tautología si y sólo si el resultado de reemplazar cualquiera de sus oraciones atómicas componentes (en todos los casos) por otras oraciones atómicas, es siempre una oración verdadera,*

* En rigor, no puede hablarse de combinaciones. Se trata, en efecto, de las ordenaciones de m objetos, tomados de n en n, con repetición, caso en que se puede tener $n > m$. El análisis combinatorio clásico da la siguiente fórmula: $O(R)^n_m = m^n$. (N. del T.)

la relación de esta definición con la prueba de la tautología constituida por la tabla de verdad debe ser clara.[1] Una línea determinada de la tabla representa la prueba de una combinación particular de oraciones atómicas. Puesto que en un análisis aleticofuncional sólo la verdad o falsedad de las oraciones atómicas produce la verdad o falsedad de la oración total, una vez que hayan sido sometidas a prueba todas las combinaciones posibles de verdad y falsedad, habrán también sido sometidas a prueba los efectos de todas las sustituciones posibles de oraciones atómicas. Como hemos visto, si una oración contiene sólo una oración atómica, hay sólo dos posibilidades

T
F

Si consta de dos oraciones atómicas distintas, hay cuatro.

T T
T F
F T
F F

Si se compone de tres oraciones atómicas diferentes, entonces son ocho:

T T T
T T F
T F T
T F F
F T T
F T F
F F T
F F F

Y, en general, como ya hemos hecho notar, si hay n oraciones atómicas componentes diferentes, hay 2^n combinaciones posibles de valores de verdad y, en consecuencia 2^n líneas en la tabla de verdad.

La frase 'en todos los casos' se añade entre paréntesis a la definición formal para dejar claramente establecido que debe recibir el mismo tratamiento una oración atómica, por cada vez que aparezca en una oración compuesta. Por ejemplo, en la tabla de verdad de $\mathbf{P} \vee \mathbf{Q} \rightarrow \mathbf{P}$ dada anteriormente, a una ocurrencia o aparición de **P** se le da el valor T cuando y sólo cuando se le da también a la otra. Este particular se aclara aún más usando el formato que sigue (que no es conveniente para fines de cálculo):

[1] Se entiende que se trata solamente de oraciones aleticofuncionales.

P	V	Q	→	P
T	T	T	T	T
T	T	F	T	T
F	T	T	F	F
F	F	F	T	F

Nótense las columnas idénticas para ambos casos u ocurrencias de **P**[2].

Es también importante advertir que si se omite la exigencia de que **P** y **Q** sean oraciones atómicas distintas, pueden encontrarse ejemplos de **P** y **Q** para los cuales **P** V **Q** → **P** es una tautología; por ejemplo, sea **P** 'llueve o no llueve' y **Q** 'hace calor'. Fácilmente se demuestra entonces por medio de la tabla de verdad derivada correspondiente que la oración 'si llueve o no llueve o hace calor, entonces llueve o no llueve' es una tautología. Además, si **P** y **Q** son la misma oración, entonces **P** V **Q** → **P** es una tautología.

Finalmente, debe también advertirse que si una oración es una tautología podemos sustituir cualquier oración atómica componente (en todas sus ocurrencias) por una oración compuesta y el resultado será una tautología. Por ejemplo, la correspondiente tabla de verdad muestra fácilmente que si **P** es una oración atómica, **P** → **P** es una tautología, pero una vez que se demuestra esto, se sigue fácilmente que **P** → **P** es una tautología cuando **P** es una oración cualquiera.

Los dos últimos párrafos pueden resumirse en dos reglas útiles:

(I) *Un enunciado que no es una tautología, puede convertirse en tautología mediante la sustitución de oraciones atómicas por oraciones compuestas o mediante la sustitución de oraciones atómicas distintas por la misma oración atómica, en el enunciado original.*[3]

(II) *Un enunciado que es una tautología sigue siéndolo cuando sus oraciones atómicas componentes, en todas sus ocurrencias, se sustituyen por oraciones en cualesquiera.*

§1.7. Implicación y equivalencia tautológicas. Se dice que una oración **P** *implica tautológicamente* una oración **Q**, si y sólo si la oración condicional **P** → **Q** es una tautología. Por ejemplo, la oración 'Locke fue soltero y Newton nunca se casó' implica tautológi-

[2] NOTA HISTÓRICA. La prueba tabular de la tautología se debe esencialmente al lógico y filósofo americano Charles S. Peirce (1839-1914). El uso en lógica del término "tautología" es un tanto reciente, pues fue propuesto por Ludwig Wittgenstein en su libro *Tractatus Logico-Philosophicus*, Londres, 1922. Algunos filósofos llaman tautología a cualquier verdad lógica o matemática, pero en este libro no se necesita esa acepción más amplia y un tanto oscura.

[3] Debe ser obvio que no toda oración puede convertirse en tautología mediante esa sustitución. De hecho, no puede una oración que es la negación de una tautología.

camente 'Newton nunca se casó', pues para dos oraciones cualesquiera $\mathbf{P}_1$ y $\mathbf{P}_2$ sabemos que la oración $\mathbf{P}_1 \,\&\, \mathbf{P}_2 \rightarrow \mathbf{P}_2$ es una tautología. Este ejemplo puede dar lugar a la falsa idea de que la noción de implicación tautológica es trivial. Como veremos en el capítulo siguiente, es de hecho fundamental para la teoría de la inferencia lógica. Si **P** implica tautológicamente a **Q**, entonces cuando **P** es verdadera **Q** debe ser verdadera. Nunca puede suceder que **P** sea verdadera y **Q** falsa, pues se requiere que $\mathbf{P} \rightarrow \mathbf{Q}$ sea una tautología. Así pues, de la premisa **P** podemos con certeza inferir **Q** sin recurrir a ningunas otras premisas.

Mediante el uso de tablas de verdad es fácil determinar si una oración implica tautológicamente a otra. La prueba es simplemente: cualquier línea que tiene la anotación 'T' para la primera oración debe tener la misma anotación para la segunda. Podemos considerar el ejemplo anterior, esto es, $\mathbf{P}_1 \,\&\, \mathbf{P}_2$ implica tautológicamente a $\mathbf{P}_2$.

P_1	P_2	$P_1 \,\&\, P_2$ *
T	T	T
T	F	F
F	T	F
F	F	F

Notamos que la oración $\mathbf{P}_1 \,\&\, \mathbf{P}_2$ tiene la anotación 'T' solamente en la primera fila y $\mathbf{P}_2$ tiene la anotación 'T' también en esta línea. Por otra parte, $\mathbf{P}_2$ no implica tautológicamente a $\mathbf{P}_1 \,\&\, \mathbf{P}_2$, pues $\mathbf{P}_2$ tiene en la tercera línea la anotación 'T', mientras que $\mathbf{P}_1$ y $\mathbf{P}_2$ tiene la entrada 'F'.

Cuando dos oraciones se implican tautológicamente en forma mutua, se dice que son *tautológicamente equivalentes*. La noción de equivalencia tautológica es más intensa que la de implicación tautológica; su papel en la inferencia no es tan central como el de la implicación tautológica, pero es importante. La razón de esta importancia no es difícil de determinar. Si dos oraciones son tautoló-

* Si el autor hubiera agregado una columna más a esta tabla de verdad, la tautología hubiera resultado evidente:

P_1	P_2	$P_1 \,\&\, P_2$	$P_1 \,\&\, P_2 \rightarrow P_2$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	T

pues la tabla anterior cumple con el requisito establecido por el propio autor de que la última columna contenga únicamente símbolos de verdad (T). (N. del T.)

gicamente equivalentes, expresan esencialmente los mismos hechos y, en consecuencia, sus papeles en la inferencia son casi idénticos.

Por vía de ejemplo, sea **A** 'Aristóteles era zurdo' y sea **L** 'Leibniz era zurdo'. Entonces la oración **A** & –**L** es tautológicamente equivalente a la oración –(–**A** V **L**). Para ver esto, podemos usar tablas de verdad.

A	L	–A	–L	–A ∨ L	–(–A ∨ L)	A & –L
T	T	F	F	T	F	F
T	F	F	T	F	T	T
F	T	T	F	T	F	F
F	F	T	T	T	F	F

La prueba es clara: las columnas que corresponden a las dos oraciones deben coincidir línea a línea en sus anotaciones, a fin de que las dos oraciones sean tautológicamente equivalentes. Esta prueba queda satisfecha por las columnas 6 y 7 de la tabla anterior y concluimos que –(–**A** V **L**) es tautológicamente equivalente a **A** & –**L**.

Quizás valga la pena hacer notar que **P** y **Q** son tautológicamente equivalentes cuando y sólo cuando la bicondicional **P** ↔ **Q** es una tautología.

EJERCICIOS

1. Un ejemplo claro de conectivo no aleticofuncional es el de posibilidad. Por ejemplo, la oración.

(1) Es posible que haya vida en Marte

es verdadera para cualquier libre interpretación de la noción de posibilidad; pero entonces también lo es la oración:

(2) Es posible que no haya vida en Marte.

Por otra parte, la oración:

(3) Es posible que 2 + 2 = 5.

es comúnmente considerada como falsa. Usando el símbolo '◇' para 'es posible que', **M** para 'hay vida en Marte' y **W** para '2 + 2 = 5', obtenemos el siguiente análisis tabular de (1)—(3):

M	–M	W	◇M	◇–M	◇W
T	F	F	T	T	F
F	T		T	T	

El análisis de $\Diamond \mathbf{M}$ y $\Diamond - \mathbf{M}$ trae como consecuencia que el único análisis aleticoorgánico del conectivo de posibilidad es que para cualquier oración **P**, $\Diamond \mathbf{P}$ es verdadera, pero el valor de verdad de $\Diamond \mathbf{W}$ controvierte esto; y vemos que no existe un adecuado análisis aleticoorgánico.

Dé ejemplos y un análisis para demostrar que los siguientes conectivos no son aleticoorgánicos:

(a) 'El Sr. Smith cree que...'
(b) 'Es necesario que...'

2. ¿Cuál de los conectivos aleticoorgánicos introducidos en este capítulo es un sinónimo aproximado del conectivo 'a menos que'? (Sugestión: Decir habrá paz, a menos que haya una guerra importante dentro de los cinco años siguientes' equivale a decir 'Si no hay una guerra importante dentro de los próximos cinco años, habrá paz').

3. Traduzca las siguientes oraciones compuestas a notación simbólica y use letras en vez de oraciones atómicas.

(a) O el fuego fue producido por incendio premeditado o fue producido por combustión espontánea.
(b) Si el agua es clara, entonces Enrique puede ver el fondo de la alberca, o es un bobo.
(c) O Juan no está aquí o María está, y Elena ciertamente está.
(d) Si hay más gatos que perros, entonces hay más caballos que perros y hay menos serpientes que gatos.
(e) El hombre de la luna es una patraña y si lo mismo es verdad de Santa Claus, muchos niños son engañados.
(f) Si las pelirrojas son amables y las rubias no tienen pecas, entonces la lógica es confusa.
(g) Si las habitaciones son escasas o la gente quiere vivir con sus parientes políticos y si la gente no quiere vivir con sus parientes políticos, entonces las habitaciones son escasas.
(h) Si Juan rinde testimonio y dice la verdad, será encontrado culpable; y si no rinde testimonio, será encontrado culpable.
(i) O Juan debe rendir testimonio y decir la verdad, o no tiene que rendir testimonio.

4. En los siguientes ejemplos determínese el valor de verdad de las oraciones componentes, por medio de los valores de verdad dados de las oraciones componentes (i)—(iv).

(i) 'Galileo nació antes que Descartes' es verdadera.
(ii) 'Descartes nació en el siglo dieciséis' es verdadera.
(iii) 'Newton nació antes que Shakespeare' es falsa.
(iv) 'Racine era un compatriota de Galileo' es falsa.

(a) Si Galileo nació antes que Descartes, entonces Newton no nació antes que Shakespeare.
(b) Si Racine era un compatriota de Galileo o Newton nació antes que Shakespeare, entonces Descartes nació en el siglo dieciséis.
(c) Si Racine no era un compatriota de Galileo, entonces o Descartes no nació en el siglo dieciséis o Newton no nació antes que Shakespeare.

5. Sean.

N = 'Nueva York es más grande que Chicago';
W = 'Nueva York está al norte de Washington';
C = 'Chicago es más grande que Nueva York'.
(**N** y **W** son verdaderas y **C** es falsa).

¿Cuáles de las oraciones siguientes son verdaderas?

(a) $N \vee C$
(b) $N \& C$
(c) $-N \& -C$
(d) $N \leftrightarrow -W \vee C$
(e) $W \vee -C \rightarrow N$
(f) $(W \vee N) \rightarrow (W \rightarrow -C)$
(g) $(W \leftrightarrow -N) \leftrightarrow (N \leftrightarrow C)$
(h) $(W \rightarrow N) \rightarrow [(N \rightarrow -C) \rightarrow (-C \rightarrow W)]$

6. Sean

P = 'Jane Austen fue contemporáneo de Beethoven';
Q = 'Beethoven fue contemporáneo de Gauss';
R = 'Gauss fue contemporáneo de Napoleón';
S = 'Napoleón fue contemporáneo de Julio César.'
(**P**, **Q** y **R** son verdaderas y **S** es falsa).

Determínense los valores de verdad de las siguientes oraciones:

(a) $(P \& Q) \& R$
(b) $P \& (Q \& R)$
(c) $S \rightarrow P$
(d) $P \rightarrow S$
(e) $(P \& Q) \& (R \& S)$
(f) $P \& Q \leftrightarrow R \& -S$
(g) $(P \leftrightarrow Q) \rightarrow (S \leftrightarrow R)$
(h) $(-P \rightarrow Q) \rightarrow (S \rightarrow R)$
(i) $(P \rightarrow -Q) \rightarrow (S \leftrightarrow R)$
(j) $(P \rightarrow Q) \rightarrow [(Q \rightarrow R) \rightarrow (R \rightarrow S)]$
(k) $P \rightarrow [Q \leftrightarrow (R \rightarrow S)]$

7. Sea **P** una oración tal, que para cualquier oración **Q** es verdadera la oración **P** V **Q**. ¿Qué puede decirse sobre el valor de verdad de **P**?

8. Sea **P** una oración tal que para cualquier oración **Q** la oración **P** y **Q** es falsa. ¿Qué puede decirse del valor de verdad de **P**?

9. Si **P** $\leftrightarrow$ **Q** es verdadera, ¿qué puede decirse sobre el valor de verdad de **P** V **–Q**?

10. Sean **P**, **Q** y **R** tres oraciones atómicas distintas. Determínese por medio de tablas de verdad cuáles de las oraciones siguientes son tautologías:

(a) $P \vee Q$
(b) $P \vee -P$
(c) $P \vee Q \rightarrow Q \vee P$
(d) $P \rightarrow (P \vee Q) \vee R$
(e) $P \rightarrow (-P \rightarrow Q)$
(f) $(P \rightarrow Q) \rightarrow (Q \rightarrow P)$
(g) $[(P \rightarrow Q) \leftrightarrow Q] \rightarrow P$
(h) $P \rightarrow [Q \rightarrow (Q \rightarrow P)]$
(i) $P \& Q \rightarrow P \vee R$
(j) $[P \vee (-P \& Q)] \vee (-P \& -Q)$
(k) $P \& Q \rightarrow (P \leftrightarrow Q \vee R)$
(l) $[P \& Q \rightarrow (P \& -P \rightarrow Q \vee -Q)] \& (Q \rightarrow Q)$

11. Si **P** y **Q** son oraciones atómicas distintas, ¿cuáles de las siguientes son tautologías?

(a) $P \leftrightarrow Q$
(b) $P \leftrightarrow P \vee P$
(c) $P \vee Q \leftrightarrow Q \vee P$
(d) $(P \rightarrow Q) \leftrightarrow (Q \rightarrow P)$
(e) $(P \leftrightarrow P) \leftrightarrow P$

12. Con base en su significado usual constrúyanse tablas de verdad para las conectivos oracionales usados en los ejemplos siguientes:

(a) No ambos **P** y **Q**. (b) Ni **P** ni **Q**.

13. Dé ejemplos de oraciones (no necesariamente atómicas) tales que resulten tautologías las siguientes oraciones compuestas:

(a) $P \& Q$
(b) $P \vee (P \& -Q)$
(c) $P \rightarrow P \& -Q$
(d) $P \rightarrow -P$

14. ¿Existe una oración **P** tal que **P** & —**P** sea una tautología?

15. Si **P** y **Q** son oraciones atómicas distintas, la oración —**P** V **Q** ¿cuál implica tautológicamente de las siguientes?

(a) P
(b) Q
(c) $P \vee Q$
(d) $P \& -Q$
(e) $-P \vee Q$
(f) $-Q \rightarrow P$
(g) $P \leftrightarrow Q$

16. Si **P** y **Q** son oraciones atómicas diferentes, ¿a cuál implica tautológicamente de las siguientes la oración —**P** V **Q**?

(a) P
(b) $Q \rightarrow P$
(c) $P \rightarrow Q$
(d) $-Q \rightarrow -P$
(e) $-P \& Q$

17. Si **P** y **Q** son oraciones atómicas distintas, ¿a cuál de las siguientes es tautológicamente equivalente la oración **P**?

(a) $P \vee Q$
(b) $P \vee -P$
(c) $P \& P$
(d) $P \rightarrow P$
(e) $-P \rightarrow P$
(f) $P \rightarrow -P$
(g) $Q \vee -Q \rightarrow P$

CAPITULO 2

TEORIA ORACIONAL DE LA INFERENCIA

§2.1. Dos criterios principales de la inferencia e interpretaciones oracionales. En este capítulo trataremos la teoría de la inferencia lógica. Las reglas de inferencia que rigen los conectivos oracionales resultan ser completamente sencillas. Puede encontrarse útil considerar las reglas presentadas en este capítulo y el No. 4 como la exposición detallada de la forma de jugar un juego no muy complicado. El juego tiene la siguiente forma: comenzamos con un conjunto de fórmulas que llamamos *premisas*. El objeto del juego consiste en aplicar las reglas de manera que se obtenga alguna otra fórmula *dada* (la conclusión deseada). El conjunto de premisas corresponde a la posición inicial de un jugador en un juego. Por una sucesión de jugadas, en que cada jugada está sancionada por una regla, llegamos a la posición del triunfo: la conclusión buscada. Como en un juego, las reglas permiten toda clase de jugadas tontas; el problema radica en aprender a ejecutar las jugadas pertinentes. (Al aprender la teoría de la inferencia será útil recordar que ordinariamente aprendemos un juego por medio de ejemplos e ilustraciones y recurrimos al enunciado formal de las reglas sólo para dilucidar puntos controvertidos o para eliminar confusiones).

Ahora bien, en un juego como el bridge o el ajedrez se escogen reglas que se supone han de dar como resultado algo interesante o entretenido. En cambio, la teoría de la inferencia lógica es algo más que entretenida. Son muchas las consideraciones que conducen a

* El capítulo anterior y el presente contienen lo que en lógica simbólica se conoce como cálculo proposicional. Como en inglés existe el tecnicismo lógico "proposition", con el mismo significado que el vocablo castellano "proposición" y de hecho el autor no lo usa, sino exclusivamente la palabra inglesa "sentence", nos ha parecido una mejor interpretación de su pensamiento el uso exclusivo de las palabras castellanas "oración" y "oracional". Esta opinión se ve confirmada por el propio autor, quien más adelante, en la Pág. 86 justifica el uso del término oración (sentence) por el hecho de atribuirle un significado especial. (N. del T.)

la construcción de un conjunto de reglas de inferencia y ciertos aspectos del problema de establecer ese conjunto de reglas son demasiado técnicos para discutirlos aquí. Sin embargo, hay dos criterios principales de construcción que predominan sobre todos los demás.

> CRITERIO I. *Dado un conjunto de premisas, las reglas de deducción lógica deben permitirnos inferir* SOLAMENTE *las conclusiones que lógicamente se siguen de las premisas.*
>
> CRITERIO II. *Dado un conjunto de premisas, las reglas de deducción lógica deben permitirnos inferir* TODAS *las conclusiones que lógicamente se siguen de las premisas.*

Pero, procede hacer la pregunta siguiente. ¿cómo pueden enunciarse criterios para establecer reglas adecuadas de inferencia, en función de una conclusión que *se siga lógicamente* de las premisas? ¿No consiste precisamente lo esencial de las reglas de inferencia en caracterizar explícitamente la noción de consecuencia lógica? Para evitar el cargo de círculo vicioso, es necesario dar una definición intuitivamente plausible de consecuencia lógica o de validez lógica. Una definición enteramente precisa es un tanto técnica; la idea básica, por otra parte, no es demasiado complicada, e introduce la noción de *interpretación* que resultará útil por varios conceptos. La idea consiste en que **Q** se sigue lógicamente de **P** cuando **Q** es verdadera en cada interpretación o modelo para el cual **P** es verdadera. Para nuestros fines presentes podemos definir como sigue la noción restringida de una *interpretación oracional:*

> *Una oración* **P** *es una interpretación oracional de una oración* **Q** *si sólo si* **P** *puede obtenerse de* **Q**, *reemplazando las oraciones atómicas componentes de* **Q** *por otras (no necesariamente diferentes).*

Si esta definición suena un poco rara, un ejemplo mostrará qué sencilla es. Sea $\mathbf{Q}_1$

> 'Si el sol está brillando, Mariana es feliz'.

y sea $\mathbf{P}_1$

> 'Si el batallón avanza demasiado aprisa, o el general está equivocado, la batalla está perdida'.

Entonces $\mathbf{P}_1$ es una interpretación oracional de $\mathbf{Q}_1$, pues 'el sol está brillando', es reemplazado por 'el batallón avanza demasiado aprisa o el general está equivocado' y 'Mariana es feliz' es reemplazada por 'la batalla está perdida'. Para ser una interpretación oracional de la oración particular $\mathbf{Q}_1$ $\mathbf{P}_1$ necesita tener solamente una carac-

terística: su principal conectivo oracional debe ser una implicación. En otras palabras, la interpretación oracional de una oración debe conservar su *forma oracional*. Por esta razón **Q_1** no es una interpretación oracional de **P_1**, pues la disyunción en el antecedente de **P_1** falta en **Q_1**. Por supuesto, la interpretación oracional puede tener mas estructuras, (por ejemplo, **P_1** tiene mas estructura que **Q_1**), pero no debe tener menos.

Debe entenderse que si una oración atómica componente aparece más de una vez en una oración, cualquier interpretación oracional de esa oración debe reemplazar la oración atómica componente por la misma cosa, en cada una de sus manifestaciones. Así, si **Q_2** es

> 'Si el sol está brillando, entonces llueve o el sol está brillando' y **P_2** es
>
> 'Si 1 + 2 = 4, entonces la nieve es negra, o yo soy un tonto',

en ese caso **P_2** no es una interpretación oracional de **Q_2**, pues la primera incidencia de 'el sol está brillando' en **Q_2** es reemplazada por '1 + 2 = 4', y la segunda por una oración diferente.

Ahora es ya posible establecer una condición suficiente (aunque no necesaria) para que una oración sea consecuencia lógica de otra.

(I) **Q** *se sigue lógicamente de* **P** *si son verdaderas todas las interpretaciones oracionales de la implicación* **P → Q**.[1]

(Como veremos en el Cap. 4, obtenemos una caracterización completa de la consecuencia lógica omitiendo la restricción a interpretaciones *oracionales*). Podemos relacionar (I) con **P** implicando tautológicamente **Q** por medio del argumento siguiente. Primero, llamamos *atómica* a una interpretación oracional si la interpretación consiste en reemplazar oraciones atómicas por oraciones atómicas. La definición de tautologías dada en el Cap. 1 puede entonces redactarse como sigue:

(II) *Una tautología es una oración cuyas interpretaciones oracionales atómicas son todas verdaderas.*

Y como implícitamente se ha observado en el Cap. 1, no es difícil ver que

(III) *Si es verdadera toda interpretación atómica oracional*

---

[1] El uso "si" sin "solamente si" indica que se da una condición suficiente, pero que no se enuncia nada con respecto de la necesidad. Compárese el § 1.3 para discusión del valor idiomático de necesidad y suficiencia.

de una oración, entonces es verdadera toda interpretación oracional de esa oración.

Mediante (I)-(III) llegamos al criterio operante de que **Q** se sigue lógicamente de **P** si **P** → **Q** es una tautología, o, en otras palabras:

*(IV) **Q** se sigue lógicamente de **P** si **Q** está tautológicamente implicada en **P**.*

Toda la teoría oracional de la inferencia está resumida en (IV). Para determinar si en virtud de las propiedades lógicas de los conectivos oracionales una conclusión dada se sigue lógicamente de un conjunto de premisas, necesitamos solamente construir la tabla de verdad correspondiente y ver si las premisas implican tautológicamente la conclusión. Por ejemplo, supóngase que deseamos saber si la conclusión 'Pedro va a llorar' se sigue lógicamente de las dos premisas:

(1) O María da a Pedro su juguete o Pedro va a llorar.
(2) María no da a Pedro su juguete.

Sea **P** 'Pedro va a llorar' y sea **M** 'María da a Pedro su juguete'. La argumentación puede entonces simbolizarse como sigue:

Premisa 1:	**M** ∨ **P**
Premisa 2.	−**M**
Conclusión:	**P**

Para decidir si la argumentación es válida, necesitamos determinar si la conjunción (**M** ∨ **P**) & −**M** implica tautológicamente a **P**. Para esta tarea, construimos la siguiente tabla de verdad de cuatro líneas:*

M	P	M ∨ P	−M	(M ∨ P) & −M
T	T	T	F	F
T	F	T	F	F
F	T	T	T	T
F	F	F	T	F

* Esta tabla, con una columna más, hubiera quedado como sigue:

M	P	M V P	−M	(M V P) & −M	
T	T	T	F	F	T
T	F	T	F	F	T
F	T	T	T	T	T
F	F	F	T	F	T

y de este modo la columna final hubiera claramente indicado que, en efecto, la implicación es tautológica. (N. del T.)

Y observamos que en el único caso en que las premisas son verdaderas en su conjunto (tercera línea, última columna), la conclusión **P** es también verdadera (tercera línea, segunda columna). Concluimos que el argumento es válido, esto es, que la conclusión es una consecuencia lógica de las premisas.

Sin embargo, hay dos buenas razones para desarrollar una teoría de la inferencia oracional que no consista simplemente en construir una tabla compacta de verdad para comprobar la validez de una argumentación. En primer lugar, si un conjunto de premisas y la conclusión deseada contienen cinco o más oraciones atómicas distintas, la correspondiente tabla de verdad debe tener cuando menos 32 líneas ($2^5 = 32$). No sólo es tedioso construir una tabla de verdad con 32 líneas, sino también difícil evitar cometer errores cuando está involucrado tan gran número de cálculos elementales. En segundo lugar, el método de tabla directa de verdad es adecuado para un número lastimosamente pequeño de argumentos lógicamente válidos, como veremos en el Cap. 4. Es, en consecuencia, deseable desarrollar una teoría oracional de la inferencia que tenga rápida aplicación en contextos más generales. Antes de dedicarnos a esta tarea en el apartado siguiente, procede hacer algunas observaciones generales sobre los dos criterios introducidos al comienzo de este apartado.

El criterio I dice que las reglas de inferencia deben ser *sólidas o legítimas*, esto es, que no deben permitir una inferencia falaz.

Combinando este criterio con (I), podemos obtener un criterio eficaz para probar la validez de una regla de inferencia propuesta. Primeramente vemos que **Q** debe poderse deducir de **P**, mediante la aplicación de las reglas, solamente si son verdaderas todas las interpretaciones de $\mathbf{P} \rightarrow \mathbf{Q}$. El criterio de comprobación es, pues:

> *Si una nueva regla de inferencia propuesta permite la deducción de una conclusión falsa, de premisas verdaderas, debe rechazarse.*

Esto es, si **Q** puede deducirse de **P** mediante la aplicación de la regla,. y es falsa alguna interpretación de $\mathbf{P} \rightarrow \mathbf{Q}$, debe rechazarse la regla.

Por ejemplo, supóngase que alguien propone la regla:

(A) De una oración **Q** y una oración $\mathbf{P} \rightarrow \mathbf{Q}$, podemos inferir **P**.

Entonces construimos inmediatamente la clase de contraejemplo requerido por el criterio de comprobación. Sea

P = 'Lincoln nació en Illinois'
Q = 'Lincoln nació en los Estados Unidos'.

Tanto **Q** como **P** → **Q** son claramente verdaderas, pero **P** es falsa, pues Lincoln nació en Kentucky. Mediante la aplicación de (A) hemos deducido una conclusión falsa de premisas verdaderas y no nos queda más que rechazarla.

El Criterio II dice que las reglas de inferencia deben ser *completas*, esto es, deben permitir la deducción de todas las conclusiones válidas.[2]

Las aplicaciones del Criterio II son más complicadas, pero un sencillo ejemplo puede mostrar que el criterio I sólo no es suficiente.

Supóngase que algún bobo propone como única regla de inferencia:

(B) De cualquier oración **P** podemos inferir **P**.

Está claro que (B) satisface el Criterio I, pues si partimos de la premisa verdadera **P** podemos solamente deducir **P** mismo. Pero sostener que (B) es suficiente para toda inferencia lógica viola el criterio II. Mediante la aplicación de (B) no podemos, por ejemplo, inferir la conclusión válida **P** de las premisas **M** ∨ **P** y −**M** (el ejemplo discutido anteriormente).

Haremos varias veces referencia a los Criterios I y II al discutir la teoría y la práctica de la inferencia válida. Especialmente en el Cap. 4 usaremos el Criterio I para justificar varias restricciones de las reglas de inferencia.

EJERCICIOS

1. Constrúyanse contra ejemplos específicos de las dos reglas falaces de inferencia siguientes [dése para (a) un ejemplo distinto al del texto]. Se indican los nombres tradicionales de estas dos falacias.

(a) Falacia de afirmar el consecuente: De **Q** y **P** → **Q**, podemos deducir **P**.
(b) Falacia de negar el antecedente: de −**P** y **P** → **Q**, podemos deducir −**Q**.

2. Usando el Criterio I, decídase cuáles de las reglas siguientes son reglas válidas de inferencia. Para las que se crea que son válidas, constrúyase un contraejemplo específico, para demostrar que violan el Criterio I; esto es, dése un ejemplo en que se obtenga una falsa conclusión de premisas verdaderas, mediante la aplicación de la regla no válida.

(a) De **P** y **Q**, podemos deducir **P** & **Q**.
(b) De **P** y **P** ∨ **Q**, podemos deducir **Q**.

[2] Aunque el conjunto de reglas introducido en este capítulo y el cuarto es sólido y completo, queda fuera del alcance de la Parte I establecer esto. La noción de completo se discute más pormenorizadamente en el apartado § 4.2.

(c) De $-Q$ y $P \vee Q$, podemos deducir P.
(d) De $-P$ y $P \vee Q$, podemos deducir $-Q$.

3. Supóngase que alguien propuso como única regla de inferencia: de P podemos deducir $Q \vee -Q$. ¿Viola esta regla el criterio I? ¿Viola el Criterio II? Si es así, explíquese por qué.

4. Constrúyase una regla (no válida) de inferencia que satisfaga el Criterio II, pero que viole el I.

5. Utilizando la discusión de las interpretaciones oracionales, explíquese exactamente por qué la siguiente regla de inferencia viola el Criterio I.

De 'Thomas Jefferson fue presidente' podemos inferir 'Jane Austen escribió Emma'.

(Nótese que no es este un caso de premisa verdadera y conclusión falsa, pues ambos enunciados son verdaderos).

6. ¿Cuál de las siguientes oraciones son interpretaciones oracionales de la oración 'Está lloviendo o está nevando'?

(a) Si está lloviendo, entonces está nevando.
(b) Está nevando.
(c) Está nevando o está nevando.
(d) Está nevando o está lloviendo.
(e) O está nevando o está lloviendo o no está lloviendo.

7. ¿Cuáles de las siguientes aserciones son verdaderas? Si es falsa, dé un contraejemplo.

(a) Una interpretación oracional de una oración P debe tener el mismo número de oraciones atómicas distintas que P.
(b) Una interpretación oracional de una oración no atómica P debe tener el mismo conectivo oracional principal que P.
(c) Una interpretación oracional de una oración P debe tener el mismo número de ocurrencias o presentaciones de conectivos oracionales que P.
(d) Una interpretación oracional de una oración P debe tener cuando menos tantas ocurrencias o incidencias de sentencias atómicas (no necesariamente distintas) como P.

§2.2. Las tres reglas oracionales de la deducción. Por las razones dadas en la sección anterior, reemplazamos la construcción de una tabla única de verdad para probar la validez de un argumento, por tres reglas de deducción. Por el momento trataremos solamente de dos de estas reglas. Una nos permite introducir premisas cuando se necesitan y la otra el uso paulatino de implicaciones tautológicas.

Antes de enunciar en forma precisa estas dos reglas de inferencia, podemos considerar un ejemplo para ver cómo se usan.

EJEMPLO 1. *Si no hay subsidios del gobierno para la agricultura, entonces hay controles gubernativos sobre la agricultura. Si hay controles gubernativos sobre la agricultura, entonces no hay depresión agrícola. Hay depresión o sobreproducción*

agrícolas. Es un hecho que no hay sobreproducción. Entonces hay subsidios del gobierno para la agricultura.

Deseamos deducir la conclusión 'Hay subsidios del gobierno para la agricultura' de las cuatro premisas dadas. Para mayor claridad, aquí y en lo subsecuente procederemos a la representación simbólica del argumento; el significado de los diversos números de la izquierda se explica a continuación.

1	(1) $-S \to C$	Premisa
2	(2) $C \to -D$	Premisa
3	(3) $D \vee O$	Premisa
4	(4) $-O$	Premisa
3, 4	(5) D	(3) & (4) implican tautológicamente a (5)
2, 3, 4	(6) $-C$	(2) & (5) implican tautológicamente a (6)
1, 2, 3, 4	(7) S	(1) & (6) implican tautológicamente a (7)

El significado de las letras debe ser obvio: **S** es la oración 'Hay subsidios del gobierno para la agricultura'; **C** es la oración 'Hay controles del gobierno sobre la agricultura'; **D** es la oración 'Hay depresión agrícola'; **O** es la oración 'Hay sobreproducción'. La deducción consta de siete líneas. La introducción de cada línea puede justificarse por una de las dos reglas. Las primeras cuatro líneas son justamente las premisas del argumento; las últimas tres líneas se obtienen demostrando que son implicaciones tautológicas de las precedentes. En el caso de la línea (6), por ejemplo, es fácil ver que la conjunción de las líneas (2) y (5), esto es, la conjunción

$$(C \to -D) \ \& \ D$$

implica tautológicamente

$$-C$$

En el caso de la línea (7), la conjunción de las líneas (1) y (6), esto es, la conjunción

$$(-S \to C) \ \& -C$$

implica

$$S$$

mediante el uso de una tautología similar a la que nos permite inferir la línea (6) de (2) y (5).

Nótese qué sencillas son las implicaciones tautológicas aquí utilizadas, en comparación con la implicación total necesaria para establecer de un golpe la conclusión:

$$[(-S \rightarrow C) \ \& \ (C \rightarrow -D) \ \& \ (D \vee O) \ \& \ -O] \rightarrow S.$$

Para probar esta implicación, se hubiera necesitado una tediosa tabla de verdad de dieciséis líneas. Las tres implicaciones tautológicas usadas no solamente son sencillas, sino que se encuentran a cada paso. Para dividir en sus partes constitutivas la mayor parte de las inferencias que se encuentran en la práctica, basta con diez o doce tipos sencillos de implicaciones tautológicas. (La sección siguiente se ocupará de estas tautologías útiles.)

No hemos explicado aún el significado de los números que aparecen en la columna de la izquierda, en la deducción anterior. Para cada línea de la derivación, la lista de números de la izquierda corresponde a las *premisas* de que depende esa línea. Así, la línea (1) depende solamente de sí misma, la primera premisa; de modo semejante, la línea (2) depende sólo de sí misma, la segunda premisa. Pero la línea (5) depende las premisas tercera y cuarta; la línea (6) se deduce de las líneas segunda y quinta y, en consecuencia, depende de las premisas de que dependen las líneas (2) y (5), esto es, las premisas segunda, tercera y cuarta; finalmente, la línea (7) se deduce de las líneas (1) y (6) y, por lo mismo, depende de las cuatro premisas. Así pues, se usaron dos premisas para deducir la línea (5); tres para deducir la línea (6), y cuatro para deducir la línea (7). Debe insistirse en la significación intuitiva de los números de la izquierda: cada línea es una *consecuencia lógica* del conjunto de premisas correspondientes a los números de la izquierda. Así, la línea (1) es trivialmente una consecuencia lógica de sí misma, y la línea (6), por ejemplo, es una consecuencia lógica de las premisas segunda, tercera y cuarta. Nótese que los números de la izquierda indican las premisas de que ha sido inferida la línea, quizás por una cadena muy complicada de inferencias. Por otra parte, los números de la derecha indican simplemente de que líneas en particular fue inferida la línea en forma inmediata, mediante la aplicación de una sola implicación tautológica.

Una advertencia sobre la forma de anotar los números de la izquierda. Cuando se sienta una premisa, en forma de una línea, se escribe a la izquierda el número de orden de *esa línea*, puesto que sólo depende de ella misma. No se anota el número correspondiente al total de premisas introducidas. Este es un importante pormenor que debe tenerse presente. Podemos aclararlo volviendo a formul-

la deducción anterior. En esta versión se incorporan también algunas abreviaturas útiles.

{1}	(1) D ∨ O	P
{2}	(2) –O	P
{1, 2}	(3) D	1, 2 T
{4}	(4) C → –D	P
{1, 2, 4}	(5) –C	3, 4 T
{6}	(6) –S → C	P
{1, 2, 4, 6}	(7) S	5, 6 T

Así pues, en la línea (4) se introduce la tercera premisa, pero el número que anotamos a la izquierda es 4, para indicar que esta línea depende sólo de sí misma; anotar, en cambio, el número 3, nos hubiera indicado que la línea (4) depende de la línea (3), lo que no es el caso. Observaciones similares se aplican a la línea (6). Hemos puesto también corchetes en los números de la izquierda. Esta notación adicional hace claro que una línea dada es una consecuencia lógica del *conjunto* de premisas correspondiente al conjunto de números aplicados a esa línea. (Es conveniente y acostumbrado describir un conjunto anotando los nombres de sus miembros, separados por comas, e incluyendo el total en corchetes. Así el conjunto cuyo miembros son los números 1, 2 y 4 se describe mediante la anotación{1, 2, 4}.)

En la anterior reformulación del Ej. 1 hemos abreviado 'premisa' con 'P' y frases como '(3) y (4) implican tautológicamente a (5)' con '3 y 4 T'. Los números '3' y '4' indican las líneas utilizadas anteriormente y la letra 'T' se refiere a la aplicación de una implicación tautológica. Las Reglas P y T pueden enunciarse sumariamente como sigue:

REGLA P. *Podemos introducir una premisa en cualquier punto de una deducción.*

REGLA T. *Podemos introducir una oración* S *en una deducción, si hay oraciones precedentes en la deducción tales que su conjunción implica tautológicamente a* S.[3]

La Regla P nos permite introducir un nuevo supuesto o premisa cada vez que lo deseamos. Esto puede parecer absurdo, pues se supondría que si se dispusiera de suficientes premisas, se podría probar absolutamente cualquier cosa. La medula de la cuestión está en que

[3] Se permite la conjunción de un número finito de oraciones, no solamente de dos. Podríamos así tener ((P → Q) & (Q → R)) & (R → S), que tautológicamente implica P → S.

el conjunto de números situado a la izquierda indica explícitamente las premisas que hemos utilizado en un argumento y cualquier argumento lógicamente correcto no es mejor o peor que las premisas en que descansa. Por supuesto, se comete un serio error lógico si se usa una premisa en un argumento *sin* reconocimiento explícito. Además, no se ha demostrado que una conclusión es consecuencia lógica de un conjunto dado de premisas, si el conjunto de números de la izquierda contiene un número que se refiere a una premisa que no está en el conjunto dado. Supóngase, por ejemplo, que alguien dijo que la deducción siguiente demuestra que de la premisa **P** ∨ **Q** podemos deducir **Q**:

{1}	(1) P ∨ Q	P
{2}	(2) −P	P (premisa adicional)
{1, 2}	(3) Q	1, 2 T

La deducción en sí es correcta, pero no demuestra que **Q** se sigue lógicamente de **P** ∨ **Q**, pues la línea (3) tiene dos números, no justamente el número 1, anotados a la izquierda y, en consecuencia, **Q** es una consecuencia lógica de dos premisas, **P** ∨ **Q** y −**P**.

La tercera y última regla de las deducciones oracionales es la de la prueba condicional, que llamamos *Regla C.P.* La idea general de esta regla es que podemos introducir una premisa **R**, por decirlo así, condicionalmente, y usarla en conjunción con las premisas originales, para deducir una conclusión **S** y a continuación aseverar que la implicación **R** → **S** se sigue de las premisas originales solamente. Debe ser intuitivamente claro que esta regla no viola el Criterio I y es, en consecuencia, aceptable. Si podemos inferir válidamente **S** de las premisas $\mathbf{P}_1$, $\mathbf{P}_2$, . . . , $\mathbf{P}_n$ y **R**, entonces podemos inferir **R** → **S** de $\mathbf{P}_1$, $\mathbf{P}_2$, . . . , $\mathbf{P}_n$. Pues supóngase que la primera inferencia fuera válida y la segunda no válida en algún caso particular. Entonces **R** → **S** tendría que ser falsa; pero podría suceder sólo si **R** fuera verdadera y **S** falsa. Sin embargo, si este fuera el caso, **S** sería falsa, las premisas $\mathbf{P}_1$, $\mathbf{P}_2$, . . . , $\mathbf{P}_n$, **R** verdaderas y la primera inferencia también inválida. Podemos, en consecuencia, aceptar:[4]

REGLA P.C. *Si podemos deducir* **S** *de* **R** *y de un conjunto de premisas, entonces podemos deducir* **R** → **S** *del conjunto de premisas solamente.*

Veamos ahora en un ejemplo cómo funciona la Regla C.P. (Hacemos notar que este ejemplo, juntamente con el anterior y cual-

[4] Alfredo Tarski demostró explícitamente por primera vez en 1929 que la regla de la prueba condicional era una regla válida de inferencia. Cuando se deriva de otras reglas de inferencia y no se toma como primitiva, se llama generalmente el *teorema de la deducción.*

quier ejemplo subsecuente, no pretende ser un argumento serio, basado en la realidad; en este libro no estamos obligados a sostener que las premisas son de hecho verdaderas).

EJEMPLO 2 (EL CIRCUITO DE LA LIGA NACIONAL). *Si los Cardenales quedan en tercer lugar, entonces los Esquivadores quedarán en segundo y los Bravos en cuarto. O los Gigantes no quedan en primer lugar, o los Cardenales quedan en tercero. De hecho los Esquivadores quedarán en segundo. En consecuencia, si los Gigantes quedan en primero, entonces los Bravos quedarán en cuarto.*

Usamos las letras 'C', 'D', etc., en la forma obvia; así, **C** es la oración 'Los Cardenales quedan en tercer lugar'.

{1}	(1) $C \rightarrow (D \rightarrow B)$	P
{2}	(2) $-G \vee C$	P
{3}	(3) D	P
{4}	(4) G	P
{2, 4}	(5) C	2, 4 T
{1, 2, 4}	(6) $D \rightarrow B$	1, 5 T
{1, 2, 3, 4}	(7) B	3, 6 T
{1, 2, 3}	(8) $G \rightarrow B$	4, 7 C.P.

En este ejemplo deseábamos obtener como conclusión la oración condicional $\mathbf{G} \rightarrow \mathbf{B}$. La deducción se facilitó grandemente al introducir la oración **G** como premisa adicional y usar esta premisa adicional para ayudar a deducir la oración **B**. Usamos a continuación la regla de la prueba condicional para aseverar que la oración $\mathbf{G} \rightarrow \mathbf{B}$ se sigue de las tres premisas dadas, que es lo que afirmamos en la línea (8). Nótese que a la derecha de la línea (8) anotamos 4 y 7, lo que indica que la línea (4) era nuestra premisa condicional y la línea (7) nuestra conclusión condicional. Al formar nuestra oración condicional tomando como antecedente la línea (4) (premisa condicional), al pasar de (7) a (8), retiramos el número 4 del conjunto {1, 2, 3, 4}, conjunto correspondiente a las premisas de que depende (7). El conjunto de premisas de que depende (8) corresponde entonces al conjunto {1, 2, 3}. Es importante advertir que:

(*a*) La Regla C.P. es, de todas las reglas que habrán de introducirse, la única que nos permite *reducir* el conjunto de números anotados a la izquierda.

(*b*) Cuando se aplica la Regla C.P. para reducir el conjunto de números anotados a la izquierda sólo puede eliminarse un número en una aplicación dada de la regla.

(c) Sólo puede tomarse como antecedente de la condición una línea que sea una premisa; en el Ej. 2 podemos tomar como antecedente la línea (4); sería incorrecto tomar como antecedente la línea (5), para obtener la siguiente línea (8):

(8′) $C \rightarrow B$,

pues entonces no sabríamos qué número debemos retirar del conjunto {1, 2, 3, 4}.

Las anteriores observaciones han tenido el propósito de mostrar cómo funciona la regla de la prueba condicional, pero nada se ha dicho sobre la *estrategia* en que se funda su aplicación. La mejor indicación general de estrategia es que cuando se desee deducir una conclusión que sea una implicación, se estudie siempre la posibilidad de usar una prueba condicional. En el Ej. 2, se desea una conclusión de la forma $\mathbf{P} \rightarrow \mathbf{Q}$, de modo que temporalmente tomamos **P** como una de las premisas.

Es deseable un ejemplo ligeramente más complejo que los primeros dos, para usarlo como referencia en la discusión de las tautologías del apartado siguiente. Nótese que la conclusión de este ejemplo es una implicación y que, por lo mismo, resulta adecuado el uso de una prueba condicional.

EJEMPLO 3. (UNA CARRERA DE CABALLOS). *Si* **A** *gana, entonces se colocarán* **B** *o* **C**. *Si se coloca* **B**, *entonces* **A** *no ganará. Si* **D** *se coloca, entonces* **C** *no se coloca. En consecuencia, si* **A** *gana,* **D** *no se coloca.*

{1}	(1) $A \rightarrow B \vee C$	P
{2}	(2) $B \rightarrow -A$	P
{3}	(3) $D \rightarrow -C$	P
{4}	(4) A	P
{1, 4}	(5) $B \vee C$	1, 4 T (Ley de la separación)
{2, 4}	(6) $-B$	2, 4 T (Leyes de doble negación y *modus tollendo tollens*)
{1, 2, 4}	(7) C	5, 6 T (*modus tollendo ponens*)
{1, 2, 3, 4}	(8) $-D$	3, 7 T (Leyes de doble negación y *modus tollendo tollens*)
{1, 2, 3}	(9) $A \rightarrow -D$	4, 8 C.P.

Puesto que la aplicación de tautologías particulares es la medula de las deducciónes oracionales, no listaremos ejercicios sino hasta

después de la sección siguiente. En el Ej. 3 se han dado los nombres de las tautologías particulares utilizadas.

Antes de tratar las implicaciones tautológicas útiles, hay que hacer dos observaciones generales sobre las deducciones oracionales. En primer lugar, ya debe ser claro qué cosa es una deducción oracional. En términos generales, es una sucesión de oraciones tales, que cada oración, o es una premisa, o está tautológicamente implicada por oraciones anteriores de la sucesión, o se obtiene de dos oraciones precedentes de la sucesión, por el procedimiento de condicionalización. Las oraciones o líneas de una deducción corresponden a los lances o jugadas sucesivas de un juego y cada jugada debe estar justificada por una de las tres reglas dadas.

En segundo lugar, aunque son considerables las virtudes de las deducciones oracionales en cuanto a establecer la validez de un argumento oracional, el método de la tediosa tabla de verdad por el que se puede optar tiene una ventaja teórica que debe mencionarse: suministra una prueba mecánica de la validez de un argumento oracional. Por una comprobación sistemática de la tabla de verdad adecuada podemos siempre decidir en un número finito de pasos si las premisas implican tautológicamente la conclusión y, en consecuencia, si ésta es oracionalmente válida.

En cambio, una deducción que consiste en una serie de pasos en que se aplican las tres reglas oracionales, solamente muestra si la conclusión es formalmente válida. Si la conclusión del argumento no es válida, teóricamente podríamos continuar interminablemente tratando de encontrar una deducción válida. En ningún punto podríamos decidir necesariamente que la conclusión era inválida, mediante un uso paulatino de nuevas premisas, implicaciones tautológicas y condicionalizaciones. Sin embargo, cuando aparece intuitivamente que una conclusión no es válida, puede ser que encontremos una interpretación oracional de la implicación cuyo antecedente es la conjunción de las premisas y cuyo consecuente es la conclusión, tal que en esta interpretación las premisas son verdaderas y la conclusión falsa. Este método no es mecánico como el de la tabla de verdad, pero tiene la ventaja de generalizar a toda la amplitud del esquema de inferencia tratado en el Cap. 4.

La aplicación del método de interpretación para demostrar que un argumento no es válido, no tiene complicaciones. Supóngase, por ejemplo, que alguien pretende que de las premisas:

(1) O Jane Austen escribió: "Northanger Abbey" o Emanuel Kant escribió "Los Fundamentos Metafísicos de la Ciencia Natural".

(2) Jane Austen escribió "Northanger Abbey",

puede inferir válidamente:

(3) Emanuel Kant escribió "Los Fundamentos Metafísicos de la Ciencia Natural".

El hecho de que todos los enunciados (1)-(3) son verdaderos no debe inducirnos a pensar que la argumentación es válida. Podemos simbolizar las premisas y la conclusión en la forma conocida.

(1′) $A \vee K$
(2′) A

(3′) K

Para demostrar que el argumento no es válido, buscamos una interpretación de **A** y de **K** tal que (1′) y (2′) sean verdaderas y (3′) sea falsa. En este caso es fácil encontrar esa interpretación. Sea **A** 'Jorge Washington fue el primer presidente de los Estados Unidos' y sea **K** 'Andrew Jackson fue el segundo presidente de los Estados Unidos'. En esta interpretación, tanto **A** ∨ **K** como **A** son verdaderas, puesto que **A** es verdadera, pero **K** es falsa. De aquí concluimos que (3) no se sigue lógicamente de (1) y (2), puesto que hay una interpretación oracional para la cual (1) y (2) son verdaderas y (3) es falsa.

En los ejercicios del final del apartado siguiente es forzoso aplicar el método de interpretación para demostrar que no son válidos algunos de los raciocinios dados.

§2.3. Algunas implicaciones tautológicas útiles. Las implicaciones tautológicas más frecuentemente usadas en las deducciones es probablemente la *Ley de la Separación*

$$P \,\&\, (P \rightarrow Q) \rightarrow Q.$$

Esta tautología corresponde a la regla de que de **P** y **P** → **Q** podemos inferir **Q**. Para cada una de las otras implicaciones tautológicas que estudiamos en este apartado hay una correspondiente regla de inferencia, que es simplemente un caso especial de la Regla T. Es, por supuesto, la regla correspondiente a la tautología que de hecho se use en el raciocinio. En el Ej. 3 del apartado anterior se usa la ley de la separación para deducir la línea (5) de la (1) y la (4):

(1) $A \rightarrow B \vee C$
(4) A

---

(5) $B \vee C$

El nombre en latín de la ley de la separación es *modus ponendo ponens*.

La tautología cuyo nombre en latín es *modus tollendo tollens* es similar a la ley de la separación. En este caso, en vez de afirmar (*ponendo*) el antecedente, podemos negando (*tollendo*) el consecuente de una implicación, negar (*tollens*) el antecedente. Así, pues,

$$-Q \,\&\, (P \rightarrow Q) \rightarrow -P.$$

El modo *tollendo tollens* se combina convenientemente con la *Ley de la Doble Negación*:

$$P \leftrightarrow --P.$$

En el Ej. 3 usamos estas dos tautologías para deducir la línea (6) de las líneas (2) y (4):

(2) $B \rightarrow -A$
(4) A

---

(6) $-B$

Podemos ampliar este uso de las dos tautologías juntas para comprender claramente el papel de cada una:

(2) $B \rightarrow -A$	
(4) A	
(4′) $--A$	4 T (Ley de la Doble Negación)
(6) $-B$	2, 4′ T (*modus tollendo tollens*)

Estas dos tautologías se usan mucho en combinación y, por lo general, no insertaremos explícitamente el paso (4′). No hay, por supuesto, necesidad formal de (4′), puesto que $\mathbf{Q} \,\&\, (\mathbf{P} \rightarrow -\mathbf{Q}) \rightarrow -\mathbf{P}$ es una tautología. Aquí se trata simplemente de hacer constar, para aplicarlas provechosamente después, cuáles son las tautologías más notorias y frecuentes, pero es importante tener presente que nada hay sagrado o especial respecto de estas tautologías aquí *citadas*, y se está en libertad de usar *cualesquiera* que se estimen convenientes. Hacer una lista de ciertas tautologías clave, es análogo a instruir al neófito en ajedrez sobre ciertas aperturas tipo, o a explicar cuándo enviar una bola o slam en el juego de bridge. Estas sugestiones tienen el propósito de ayudar al lector a desarrollar una estrategia operante.

En el Ej. 3 usamos la implicación tautológica *modus tollendo ponens* para deducir la línea (7) de las líneas (5) y (6):

(5) $B \vee C$
(6) $-B$

(7) C

El principio de esta tautología, como su nombre lo indica, es que si negamos (*tollendo*) un miembro de una disyunción, podemos entonces afirmar (*ponens*) el otro. Enunciada como regla de deducción, toma la forma:

De $-\mathbf{P}$ y $\mathbf{P} \vee \mathbf{Q}$ podemos deducir $\mathbf{Q}$.

Hemos citado todas las tautologías usados en los tres ejemplos de raciocinios de la sección o apartado anterior. Haremos una lista de otras tautologías más e ilustraremos el uso de algunas de ellas mediante ejemplos sencillos. A fin de que la lista quede completa incluiremos las cuatro tautologías ya discutidas.

TABLA DE TAUTOLOGIAS UTILES*

Implicaciones Tautológicas

Ley de la separación	$P \& (P \rightarrow Q) \rightarrow Q$
Modus tollendo tollens	$-Q \& (P \rightarrow Q) \rightarrow -P$
Modus tollendo ponens	$-P \& (P \vee Q) \rightarrow Q$
Ley de la simplificación	$P \& Q \rightarrow P$
Ley de la adjunción	$P \& Q \rightarrow P \& Q$
Ley del silogismo hipotético	$(P \rightarrow Q) \& (Q \rightarrow R) \rightarrow (P \rightarrow R)$
Ley de la exportación	$[P \& Q \rightarrow R] \rightarrow [P \rightarrow (Q \rightarrow R)]$
Ley de la importación	$[P \rightarrow (Q \rightarrow R)] \rightarrow [P \& Q \rightarrow R]$
Ley del absurdo	$[P \rightarrow Q \& -Q] \rightarrow -P$
Ley de la adición	$P \rightarrow P \vee Q$

Equivalencias Tautológicas

Ley de la doble negación	$P \leftrightarrow --P$
Ley de la contraposición	$(P \rightarrow Q) \leftrightarrow (-Q \rightarrow -P)$
Leyes de De Morgan	$-(P \& Q) \leftrightarrow -P \vee -Q$

* Debe quedar perfectamente claro que por medio de una tabla de verdad es posible demostrar el carácter tautológico de las tautologías aquí indicadas.

Por ejemplo, la tabla para la ley del silogismo hipotético, ley que tan importante papel desempeña en la lógica aristotélica, quedaría como sigue: (N. del T.)

P	Q	R	$P \rightarrow Q$	$(Q \rightarrow R)$	$P \rightarrow R$	$(P \rightarrow Q)$ & $(P \rightarrow R)$	$(P \rightarrow Q)$ & $(Q \rightarrow R) \rightarrow$ $Q \rightarrow R$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	T	F	T
T	F	F	F	T	F	F	T
F	T	T	T	T	T	T	T
F	T	F	T	F	T	F	T
F	F	T	T	T	T	T	T
F	F	F	T	T	T	T	T

	$-(P \vee Q) \leftrightarrow -P \,\&\, -Q$
Leyes conmutativas	$P \,\&\, Q \leftrightarrow Q \,\&\, P$
	$P \vee Q \leftrightarrow Q \vee P$
Ley de equivalencia para implicación y disyunción	$(P \rightarrow Q) \leftrightarrow -P \vee Q$
Ley de la negación para la implicación	$-(P \rightarrow Q) \leftrightarrow P \,\&\, -Q$
Una ley para oraciones bicondicionales	$(P \leftrightarrow Q) \leftrightarrow (P \rightarrow Q) \,\&\, (Q \rightarrow P)$
Otra ley para oraciones bicondicionales	$(P \leftrightarrow Q) \leftrightarrow (P \,\&\, Q) \vee (-P \,\&\, -Q)$

Dos tautologías más

Ley del tercero excluido	$P \vee -P$
Ley de contradicción	$-(P \,\&\, -P)$

Dada la premisa:

(1) Si Fillmore nació en Nueva York, entonces nació en los Estados Unidos,

deducimos la ley de contraposición:

(2) Si Fillmore no nació en los Estados Unidos, entonces no nació en Nueva York,

y por la ley de equivalencia por Implicación y Disyunción

(3) O Fillmore no nació en Nueva York o nació en los Estados Unidos;

y de (3) por las leyes de De Morgan, deducimos

(4) No es el caso que Fillmore haya nacido en Nueva York y no en los Estados Unidos.

Dada la premisa.

(5) Si la tienda está abierta y la venta continúa, entonces compraré un bote de remos,

derivamos por la ley de exportación:

(6) Si la tienda está abierta, entonces si la venta continúa, compraré un bote de remos,

y por la ley de contraposición obtenemos de (6)

(7) Si no se da el caso de que si la venta todavía sigue yo compré un bote de remos, entonces la tienda no está abierta.

Los anteriores ejemplos tienen un valor limitado a causa de que no están incluidos en raciocinios lógicos de cierta complejidad. Cuando usemos las tautologías de la tabla anterior en el apartado siguiente, así como el Cap. 4, las mencionaremos explícitamente.

EJERCICIOS

1. Enúnciense las reglas de deducción correspondientes a cada una de las tautologías de la tabla anterior, con excepción de los Principios de Tercero Excluido y de Contradicción. Cuando la tautología es una equivalencia y no una simple implicación, la regla se enuncia como la Ley de Doble Negación en el siguiente enunciado:

De **P** *podemos derivar — —* **P**, *y recíprocamente.*

Añadimos 'y recíprocamente' para indicar que la inferencia opera en ambos sentidos.

Para los raciocinios siguientes (Ejercicios 2 a 14) trátese de construir deducciones oracionales como las de los Ejs. 1 a 3 del apartado anterior. En algunos casos el raciocinio no es lógicamente válido; en esos casos escríbase 'N.S.C.' (no se sigue la conclusión) y désele entonces una interpretación oracional que demuestre su invalidez. Use las letras indicadas e identifique por su nombre los tautologías que use.

2. Si el mercado es perfectamente libre, entonces un solo proveedor no puede afectar las precios. Si un solo proveedor no puede afectar los precios, entonces hay un gran número de proveedores. Hay, en efecto, un gran número de proveedores. En consecuencia, el mercado es perfectamente libre. (**F, S, N**)

3. Si los precios son altos, entonces los salarios son altos. Los precios son altos. o hay control de precios. También, si hay control de precios, entonces no hay inflación. Sin embargo, hay inflación. En consecuencia, los salarios son altos (**P, W, C, I**)

4. Si se elevan los precios o los salarios, habrá inflación. Si hay inflación, entonces el Congreso debe regularla, o el pueblo sufrirá. Si el pueblo sufre, los congresistas se harán impopulares. El Congreso no regulará la inflación y los congresistas no se volverán impopulares. En consecuencia, no subirán los salarios. (**W, P, I, C, S, U**)

5. O la lógica es difícil o no les gusta a muchos estudiantes. Si las matemáticas son fáciles, entonces la lógica no es difícil. En consecuencia, si a muchos estudiantes les gusta la lógica, las matemáticas no son fáciles. (**D, L, M**)

6. Si Algernon está en la cárcel, entonces no es una molestia para su familia. Si no está en la cárcel, entonces no es una calamidad. Si no es una calamidad, entonces está en el ejército. Si está borracho, es una molestia para su familia. En consecuencia, o no está borracho, o está en el ejército. (**J, N, D, A, R**).

7. Si Algernon es una molestia, no está en la cárcel. Si está en la cárcel, entonces es una calamidad. Si es una calamidad, entonces no está en el ejército. En consecuencia, o no está en el ejército, o no es una molestia. (**N, J, D, A**)

8. O Juan y Enrique son de la misma edad, o Juan es de más edad que Enrique. Si Juan y Enrique son de la misma edad, entonces Elizabeth y Juan no son de la misma edad. Si Juan es de más edad que Enrique, entonces Juan es de más edad que María. En consecuencia, o Elizabeth y Juan no son de la misma edad o Juan es de más edad que María. (**S, O, E, M**)

9. Mariana opinaba que el coronel Brandon era demasiado viejo para casarse. Si la conducta de Mariana fuera siempre consistente con sus opiniones y si opinaba que el coronel Brandon era demasiado viejo para casarse, entonces no se casaría con el coronel Brandon. Pero Mariana se casó con el coronel

Brandon. En consecuencia, la conducta de Mariana no era siempre consistente con sus opiniones. (**B, C, M**)

10. Si es éste el mes de diciembre, entonces el mes pasado fue el de noviembre. Si el mes pasado fue el de noviembre entonces hace seis meses era el mes de junio. Si hace seis meses era el mes de junio, entonces hace once meses era el de enero. Si el mes que sigue es el de enero, entonces éste es el de diciembre. El mes pasado era el de noviembre, entonces éste es el de diciembre. (**D, N, J, A, X**)

11. Si María es una verdadera amiga, entonces Juan está diciendo la verdad. Si Juan está diciendo la verdad, entonces Elena no es una verdadera amiga. Si Elena no es una verdadera amiga, entonces no está diciendo la verdad. Si Elena no está diciendo la verdad, entonces María es una verdadera amiga. Pero si María es una verdadera amiga, entonces Elena no es una verdadera amiga. En consecuencia, Elena no está diciendo la verdad. (**M, J, H, T**)

12. Si y sólo si Roger ha celebrado el contrato *y* el contrato es legal, *y* no ha cumplido el contrato, Jones ganará la demanda judicial. Si Roger no ha aceptado la oferta de Jones, no ha celebrado el contrato. El hecho es que Roger no ha aceptado la oferta de Jones. En consecuencia, Jones no ganará la demanda judicial. (**R, L, P, J, A**)

13. Si Brown celebró el contrato, *o* si Brown recibió beneficios sustanciales de actos ejecutados por Smith, Brown *no* ganará el juicio. Si Brown revocó su oferta antes de que Smith la aceptara, Brown no aceptó el contrato. El hecho es que Brown no revocó su oferta antes de que Smith la aceptara. En consecuencia, Brown no ganará el juicio. (**C, B, W, R**)

14. Si Brown no celebró el contrato, o si Brown cumplió el contrato, Smith no ganará el juicio. Si Brown dejó de entregar las mercancías en la fecha convenida. En consecuencia, Smith ganará el juicio. (**C, P, W, D**)

§2.4. Consistencia de las premisas y pruebas indirectas. Algunas veces no estamos interesados en deducir una conclusión particular de un conjunto de premisas, sino en decidir si las premisas son consistentes. Esta es, a menudo, la meta principal de un abogado que interroga a un testigo en representación de la parte contraria. Si puede demostrar que el testimonio es inconsistente, habrá avanzado bastante en el propósito de demostrar la invalidez de la prueba presentada por ese testigo. La noción intuitiva de inconsistencia es que un conjunto de premisas es inconsistente si no pueden ser ciertas todas al mismo tiempo. Por ejemplo, si un testigo declara que

(1) Blenker estaba en Washington, D. C., la noche del asesinato y Blenker asegura que

(2) El, Blenker, no estaba en Washington, D. C., la noche del asesinato,

entonces sabemos de inmediato que alguien está mintiendo, pues (1) y (2) no pueden ser simultáneamente verdaderas, esto es, las aseveraciones de Blenker y el testigo son inconsistentes.

En muchos casos no es fácil decidir si un conjunto de premisas es consistente, simplemente "observándolas" y, en consecuencia,

es deseable desarrollar una técnica analítica para investigar la consistencia. Para comenzar, se dice que dos oraciones son *contradictorias** si una es negación de la otra; una contrádicción es una conjunción de dos oraciones contradictorias, esto es, es una conjunción de la forma $S \& -S$. Ahora bien, es fácil ver que un conjunto de premisas es consistente si puede deducirse lógicamente una contradicción, pues si las premisas pudieran ser todas simultáneamente verdaderas, podríamos construir un ejemplo que violara el criterio I, esto es, premisas verdaderas y la conclusión necesariamente falsa $S \& -S$. Nuestra técnica para investigar un conjunto de premisas consiste, pues, en deducir una contradicción. Abordamos el problema de intentar deducir una contradicción en la misma forma general en que abordamos el de derivar una conclusión dada. La diferencia esencial está en que al deducir una conclusión dada la meta final de la deducción se fija de antemano, mientras que al deducir una contradicción, es *cualquier* contradicción sin importar de qué contradicción se trate en particular.

El propósito del siguiente ejemplo es el de ilustrar cómo pueden usarse las reglas de inferencia para demostrar que es inconsistente un conjunto de premisas.

EJEMPLO 4. *Si el contrato es válido, entonces Horacio está jurídicamente obligado. Si Horacio está jurídicamente obligado, quebrará. Si el banco le presta dinero, no quebrará. De hecho, el contrato es válido y el banco le prestará dinero.*

{1}	(1) $V \rightarrow L$	P	
{2}	(2) $L \rightarrow B$	P	
{3}	(3) $M \rightarrow -B$	P	
{4}	(4) $V \& M$	P	
{4}	(5) V	4 T	(Ley de la simplificación)
{1, 2}	(6) $V \rightarrow B$	1, 2 T	(Ley del silogismo hipotético)
{1, 2, 4}	(7) B	5, 6 T	
{4}	(8) M	4 T	
{3, 4}	(9) $-B$	3, 8 T	
{1, 2, 3, 4}	(10) $B \& -B$	7, 9 T	(Ley de la adjudicación)

* Este concepto no coincide con el de la lógica Aristotélica, que exige que una proposición tenga "cualidad" y "cantidad" distintas de las de su proposición contradictoria. (N. del T.)

Nótese que en la prueba de que las cuatro premisas del Ej. 4 implican una contradicción, hemos usado tres implicaciones tautológicas no usadas anteriormente, tomadas de la lista del apartado anterior.

Una vez que se ha descrito la técnica para demostrar que es inconsistente un conjunto de premisas, es natural preguntar sobre la existencia de una técnica similar, para probar que es consistente un conjunto de premisas. El método de interpretación es el dispositivo adecuado. Si puede encontrarse una interpretación oracional verdadera de la conjunción de las premisas, entonces éstas son consistentes; pero si también pudiera deducirse de ellas una contradicción, tendríamos entonces premisas verdaderas en la interpretación y una válida conclusión falsa (la contradicción), en violación del criterio I. Como aplicación de este método considérense las premisas

> *Si la guerra está próxima, entonces el ejército ha sido movilizado. Si el ejército ha sido movilizado, entonces son altos los costos de la mano de obra. La guerra no está próxima y, no obstante, los costos de la mano de obra son altos.*

Podemos simbolizar premisas en la forma siguiente:

(1) $W \rightarrow A$

(2) $A \rightarrow L$

(3) $-W \,\&\, L$

Para demostrar que las tres premisas son consistentes, deseamos encontrar una interpretación **W**, **A** y **L** tal que conforme a esta interpretación las tres premisas son simultáneamente verdaderas. Es ésta una tarea fácil. Sean.

$$W = \text{'}2 + 2 = 5\text{'}$$
$$A = \text{'}1 + 1 = 2\text{'}$$
$$L = \text{'}2 + 2 = 4\text{'}$$

En este caso **W** es falsa, **A** y **L** son verdaderas y un simple análisis aleticofuncional demuestra que (1)-(3) son verdaderas. De aquí que las premisas dadas sean consistentes.

Debe notarse, sin embargo, que la consistencia de un conjunto de premisas cuya estructura lógica pueda expresarse por puros conectivos oracionales, puede ser determinada directamente por la prueba mecánica de la tabla de verdad. Se construye la tabla de verdad correspondiente a la conjunción de las premisas. Si *todas* y cada una de las anotaciones de la columna final es 'F', entonces las premisas son inconsistentes, pues no hay interpretación de sus oraciones atómicas correspondientes que las haga verdaderas. Si, *cuando menos*, una anotación es 'T', las premisas son consistentes, pues esta línea particular de la tabla de verdad sumnistra la base de una

interpretación verdadera. La dificultad con esta prueba, como la dificultad con la prueba mecánica de la tabla de verdad de la validez lógica, es que no puede ser usada una vez que se ha introducido el aparato adicional del Cap. 4.

Podemos usar la regla de la prueba condicional y la noción de un conjunto inconsistente de premisas, para introducir el importante método de *prueba indirecta* (llamado también prueba por contradicción y *prueba por reducción al absurdo*). El uso de argumentos indirectos es quizá bien conocido desde la geometría elemental. La técnica de esas pruebas se desarrolla como sigue:

(1) Introdúzcase como una nueva premisa la negación de la conclusión deseada.
(2) Partiendo de esta premisa, juntamente con las premisas dadas, dedúzcase una contradicción.
(3) Afírmese la conclusión deseada como inferencia lógica de la premisas.

Podemos demostrar esquemáticamente cómo estos tres pasos caen dentro del patrón o esquema de una deducción formal. Sea $\mathscr{P}$ la conjunción de las premisas y **C** la conclusión deseada

{1}	(1) $\mathscr{P}$	P
{2}	(2) -C	P
⋮	⋮	⋮
{1, 2}	(n) S & -S	Por reglas de deducción
{1}	($n+1$) -C → (S & -S)	2, n P.C.
{1}	($n+2$) C	$n+1$ T (Ley del absurdo

Para ilustrar este esquema, y para dar un ejemplo particular de una prueba indirecta, podemos volver a probar la validez de la conclusión del Ej. 3 del § 2.2. Nótese que la línea (4) es tautológicamente equivalente a la negación de la conclusión deseada **A** → **-D**. (El hecho de que **A & D** ↔ -(**A** → **-D**) sea una tautología se sigue de la Ley de la negación para la implicación.)

{1}	(1) A → B ∨ C	P
{2}	(2) B → -A	P
{3}	(3) D → -C	P
{4}	(4) A & D	P
{1, 4}	(5) B ∨ C	1, 4 T
{3, 4}	(6) -C	3, 4 T
{1, 3, 4}	(7) B	5, 6 T
{1, 2, 3, 4}	(8) -A	2, 7 T

{1, 2, 3, 4}	(9) A & -A	4, 8 T
{1, 2, 3}	(10) A & D → A & -A	4, 9 P.C.
{1, 2, 3}	(11) A → -D	10 T

En este ejemplo la línea (4) corresponde a la (2) del esquema, la línea (9) a la línea (n), la línea (10) a la línea ($n + 1$) y la línea 11 a la línea ($n + 2$). Nótese que (11) se sigue de (10) por la Ley del Absurdo y la Ley de la Negación para la implicación.

Está claro que la línea correspondiente a la línea ($n + 1$) del esquema [línea (10) del ejemplo anterior] va a tener la misma forma en toda prueba indirecta. La sugestión natural es la de introducir una nueva regla que permita la eliminación de esta línea intuitivamente redundante y que justifique el paso directo de la línea (n) a la línea ($n + 2$) del esquema [de la línea (9) a la (11) del ejemplo]. Puesto que esta nueva regla abrevia las deducciones en sólo una línea, no es de gran importancia práctica. Sin embargo, en el Cap. 5 se introduce cierto número de estas reglas "para abreviar" y algunas de ellas conducen a una considerable simplificación de las deducciones. Así pues, por el momento, lo que importa es que quede muy clara la estructura general de esta nueva regla de prueba indirecta. El hecho más significativo con respecto de la misma es que no pertenece a la misma categoría de las tres reglas ya introducidas. ¿Por qué? Porque usamos las tres reglas ya introducidas para establecer su validez lógica. En otras palabras, la regla de las pruebas indirectas es una regla *derivada* más bien que una regla original.[5]

> REGLA DERIVADA PARA PRUEBAS INDIRECTAS: R. A. A. *Si una contradicción es derivable de un conjunto de premisas más la negación de una fórmula* S, *entonces* S *es derivable del conjunto de premisas únicamente.*

Usamos la expresión 'R. A. A.' para representar '*reductio ad absurdum*'.

PRUEBA DE LA REGLA DERIVADA: R. A. A. Sea $\mathscr{P}$ el conjunto de premisas. Por hipótesis, deducimos de $\mathscr{P}$ y -T una contradicción, esto es, una oración S & -S. Por la regla de la prueba condicional establecemos a continuación que la oración -T → S & -S es deducible de $\mathscr{P}$. Usando a continuación la implicación tautológica que hemos llamado la Ley del Absurdo en el § 2.4, deducimos T.

[5] En este respecto, el lector quizá se preguntaría si alguna de las tres reglas originales podría haber sido derivada de las otras dos. Tal como están formuladas en el texto, la respuesta es negativa; cada una es independiente de las otras dos. Se vuelve a este particular en el Cap. 5.

Hemos escrito esta prueba en la forma anterior, para poner de relieve que la prueba de una regla derivada (esto es, un procedimiento abreviado) *no* es una deducción. La prueba no puede ser una deducción, pues una deducción se efectúa por medio de fórmulas definidas, mientras que la prueba de una regla derivada debe ser dada en una forma que se aplique a cualesquiera fórmulas.

El uso de la Regla R. A. A. se ilustra en el ejemplo siguiente.

> EJEMPLO 5. *Si veinticinco divisiones son suficientes, entonces el general ganará la batalla. O se suministran tres alas de apoyo aéreo táctico, o el general no ganará la batalla. También, no es cierto que sean suficientes veinticinco divisiones y que se vayan a suministrar tres alas de apoyo aéreo táctico. En consecuencia, no son suficientes veinticinco divisiones.*

{1}	(1) $D \rightarrow W$	P
{2}	(2) $A \vee -W$	P
{3}	(3) $-(D \& A)$	P
{4}	(4) D	P
{1, 4}	(5) W	1, 4 T
{1, 2, 4}	(6) A	2, 5 T
{3}	(7) $-D \vee -A$	3 T (Leyes de de Morgan)
{1, 2, 3, 4}	(8) $-D$	6, 7 T
{1, 2, 3, 4}	(9) $D \& -D$	4, 8 T
{1, 2, 3}	(10) $-D$	4, 9 R.A.A.

En este ejemplo, la Regla R.A.A. se usa para deducir la línea (10) directamente de las líneas (4) y (9). Puesto que un raciocinio por *reductio ad absurdum* es un caso especial de la regla de prueba condicional, permite una reducción en el conjunto de números adjudicados a una línea dada. Así, en el ejemplo anterior, {1, 2, 3, 4} están adjudicados a la (9) y el conjunto reducido {1, 2, 3} está adjudicado a (10). Intuitivamente esta situación refleja el hecho de que la contradicción de la conclusión, línea (4), es una premisa introducida solamente para fines de trabajo.

La decisión de intentar una prueba indirecta en vez de una directa resulta siempre determinada por cierto número de factores. Una prueba indirecta es sugerida en forma inmediata por una situación en que no parece haber suficientes premisas, pues, cuando uno se aventura a encontrar una prueba indirecta, le es permitido introducir la negación de la conclusión como otra premisa. Aunque el Ej. 5 implica una inferencia muy sencilla, ilustra este punto. La conclusión está representada por una simple negación, de modo que

no es conveniente una prueba condicional directa. Además, ninguna de las tres premisas es una oración atómica simple. Teniendo enfrente una implicación y dos disyunciones, es difícil ver cómo comenzar. Niéguese la conclusión, obténgase una oración atómica como premisa extra, y la maquinaria deductiva está lista para elaborar una respuesta.

EJERCICIOS

1. ¿Cuál de los siguientes conjuntos de premisas es inconsistente? Si un conjunto es inconsistente, deduzca una contradicción. Si es consistente, dé una interpretación oracional verdadera para probarlo

(a) Si el contrato es válido, entonces Horacio está jurídicamente obligado. Si Horacio está jurídicamente obligado, entonces quebrará. Horacio quebrará, o el banco le prestará dinero. Sin embargo, el banco definitivamente no le prestará dinero. (**V, L, B, M**)

(b) Si Jones cometió el crimen, entonces estaba en el apartamiento de la víctima y no salió antes de las once. En realidad, estaba en el apartamiento de la víctima. Si salió antes de las once, entonces el portero lo vio. Pero no es el caso que el portero lo haya visto ni que haya cometido el crimen. (**M, A, L, D**)

(c) El contrato se satisface si y sólo si el edificio se termina para noviembre 30. El edificio se termina si y sólo si el subcontratista de la instalación eléctrica termina su trabajo para noviembre 10. El banco pierde dinero si y sólo si el contrato no se cumple. No obstante, el subcontratista de la instalación eléctrica termina su trabajo para noviembre 10 si y sólo si el banco pierde dinero. (**C, B, E, L**)

(d) Smiling Pete ganará la cuarta carrera si y sólo si Crazy Cat llega en tercer lugar. Si Crazy Cat llega en tercer lugar, entonces Dapper Dan llegará el último. O Dapper Dan no llega el último, o Wildwood empata en primer lugar. Si Smiling Pete no gana la cuarta carrera, entonces Wildwood empatará en primer lugar. Además, Wildwood no empatará en primer lugar. (**S, C, D, W**)

(e) $P \rightarrow Q$
$Q \leftrightarrow R$
$R \vee S \leftrightarrow -Q$

(f) $-(-Q \vee P)$
$P \vee -R$
$Q \rightarrow R$

2. Demuestre que de un conjunto inconsistente de premisas puede derivarse cualquier fórmula.

3. Demuéstrese, de ser posible por *reducción al absurdo, que*:

(a) Si Juan juega primera base y Smith lanza contra nosotros, ganará Winsocki. O. Winsocki no ganará, o entonces la novena terminará a la cola de la liga. La novena no terminará a la cola de la liga. Además, Juan jugará primera base. En consecuencia, Smith no lanzará contra nosotros. (**J, S, W, T**)

(b) Si el hedonismo no es correcto, entonces el erotismo no es virtuoso.

Si el erotismo no es virtuoso, entonces o el deber no es la virtud más alta, o el supremo deber es la prosecución del placer. Pero el supremo deber no es la prosecución de placer. En consecuencia, o el deber no es la virtud más alta, o el hedonismo no es correcto. (**H, E, D, S**)

(c) Si una declaración de guerra es una estrategia adecuada, entonces o cinçuenta divisiones están acantonadas en la frontera, o veinte alas de bombarderos de largo alcance están listas para atacar. Pero no están acantonadas en la frontera cincuenta divisiones. En consecuencia, si no están listas para atacar veinte alas de bombarderos de largo alcance, entonces, o una declaración de guerra no es una estrategia adecuada, o hay armas secretas disponibles. (**D, F, T, S**)

4. Pruebe por *reducción al absurdo*:

(a) Ejercicio 5 del §2.3.

(b) Ejercicio 8 del §2.3.

CAPITULO 3

REPRESENTACION SIMBOLICA DEL LENGUAJE COTIDIANO

§3.1. Gramática y lógica. En los dos capítulos anteriores hemos desarrollado la lógica de los conectivos oracionales. Nuestra notación ha consistido únicamente en los símbolos '−', '&', '∨', '→', '↔', paréntesis y letras solas que simbolizan oraciones. Un momento de reflexión es suficiente para que cualquiera advierta que un aparato tan magro como éste no es suficiente para expresar distinciones lógicas de la clase más elemental. Por ejemplo, dentro de este marco sí es posible formular las sencillas verdades de la aritmética. De hecho, nuestro aparato lógico no sería adecuado para las necesidades de un niño de tres años de edad. No hay modo de simbolizar nombres propios y comunes, pronombres, verbos, adjetivos o adverbios; no es posible indicar las distinciones gramaticales más comunes. Por supuesto, no hemos adquirido el compromiso de desarrollar una lógica adecuada a todos los matices del lenguaje cotidiano. Lo que necesitamos es un número suficiente de nociones lógicas para expresar cualquier conjunto de hechos sistemáticos. Introduciremos específicamente tres clases de expresiones, correspondientes a las muchas clases que se distinguen en la gramática del lenguaje usual. Consideraremos *términos, predicados y cuantificadores* y nos ocuparemos de ver hasta qué punto podemos traducir oraciones del lenguaje cotidiano a un lenguaje que consiste únicamente en conectivos oracionales, términos, predicados, cuantificadores y paréntesis. Nos interesaremos también en comparar la gramática de nuestro simbolismo lógico con la del lenguaje cotidiano.

Debe tomarse nota, exactamente al principio de esta discusión, que el significado del lenguaje usual no es suficientemente uniforme y preciso para permitir el enunciado de reglas de transcripción, categóricas y exentas de ambigüedad. Al transcribir una oración en es-

pañol a nuestro simbolismo, es a menudo necesario tomar cuidadosamente en cuenta varias posibilidades.

§3.2. **Términos.** Se dará ahora una definición precisa de 'término'. Quizá el mejor modo de comenzar es estudiando la clase más importante de términos de la lógica y las matemáticas, a saber, las *variables*. Los estudiantes que se inician en la lógica algunas veces encuentran un tanto confusa la noción de variables. El punto importante que debe fijarse en la memoria es que las variables son simplemente letras, como 'x', 'y', 'z', o letras con subíndices, como 'x_1', 'x_2', 'x_3', ... La función gramatical de las variables es semejante a la de los *pronombres* y *nombres comunes* del lenguaje usual. Por ejemplo, la oración:

(1) Toda cosa es roja o no roja

es equivalente a la oración:

(2) Para toda x, x es roja o x es no roja.

En este ejemplo, el uso de la variable 'x' en (2) corresponde al uso de 'cosa' en (1). La variable particular 'x' no tiene importancia especial, y (2) puede ser formulada nuevamente usando 'y' en vez de 'x':

Para toda y, y es roja o y no es roja.

Desde un punto de vista literario (2) parece ser una versión un tanto bárbara de (1). Sin embargo, el uso de las variables está indicado cuando se trata aún de los más sencillos enunciados aritméticos. Por ejemplo, todos sabemos que para cualesquiera números x, y y z,

(3) $$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Una interpretación verbal de (3) es torpe y está lejos de ser clara. Quedaría más o menos como sigue:

(4) Si un número se multiplica por la suma de un segundo y un tercer número, entonces el resultado es igual a la suma del producto del número por el segundo número, más el producto del número por el tercer número.

Aún más torpe que (4) es el enunciado de un proceso de raciocinio en aritmética, sin el uso de variables.

Los nombres propios forman otra importante clase de términos. A fin de obtener una noción compacta, usamos letras minúsculas del principio del alfabeto para representar nombres propios. Así, si hacemos

$a =$ Isaac Newton

podemos traducir:

(5) Isaac Newton es el más grande matemático de los últimos tres siglos,

como:

(6) *a* es el más grande matemático de los últimos tres siglos.

Por el momento, (6) no parece ser mucho más compacto que (5), pero se permite una ulterior y sustancial abreviatura cuando notamos que descripciones definidas como 'el más grande matemático de los tres últimos siglos' funcionan exactamente como nombres propios y forman así otra clase de términos. Puesto que usamos también letras minúsculas del principio del alfabeto para representar descripciones definidas, si hacemos

$b =$ el más grande matemático de los últimos tres siglos,

podemos ahora traducir (5) como:

(7) *a* es *b*

El significado de "es" en (5) y (7) es el de identidad y podemos escribir (7):

$$a = b$$

Hacemos la observación de que los nombres propios o las descripciones de objetos reciben algunas veces la designación de 'constantes'. Por ejemplo, '*a*' y '*b*' e 'Isaac Newton' son constantes.

Hasta ahora hemos considerado tres clases de términos: variables, nombres propios y descripciones definidas. La idea intuitiva de término es que designa o nombra un ente. Por otra parte, las variables no nombran nada. Por ejemplo, en la expresión

$$x > 3$$

la variable '*x*' no designa un número único como 5 o 17. No obstante, la idea intuitiva de que los términos sirven para designar es correcta y podemos usarla para formular una definición general.

Definición. *Un término es una expresión que nombra o describe algún objeto, o resulta en un nombre o descripción de un objeto cuando las variables de la expresión son sustituidas por nombres o descripciones.*

Por ejemplo, de acuerdo con la definición, '*x*' es un término, porque cuando reemplazamos '*x*' por el número arábigo '3' obtenemos una

expresión que designa al número tres. La definición clasifica también como términos a expresiones como '$x + y$', pues cuando reemplazamos 'x' por '2' y 'y' por '3', obtenemos '$2 + 3$' expresión que designa, (esto es, nombra) al número 5. Algunos otros ejemplos de términos son:

$x + 3$
$x^2 + y^2 - 1$
el hombre más gordo entre los habitantes de x.

§3.3. Predicados. Además de nuestra notación para los conectivos oracionales y los términos, usamos letras mayúsculas para representar predicados. Así, podemos usar la letra '*R*' en vez del predicado 'es rojo' y simbolizar (2) apartado anterior como sigue:

Para cada x, $Rx \vee -Rx$

Debe quedar claramente establecido que '$-Rx$' se lee 'x no es rojo'; no implica lo anterior ningún uso nuevo de la negación, pues '$-Rx$' es simplemente la negación de la fórmula 'Rx'.

Ahora bien, en la gramática tradicional un predicado es la palabra o palabras en una oración o cláusula que expresa lo que se dice del sujeto. En español es posible formar predicados de varios modos: un verbo sólo, un verbo con un adverbio o locución adverbial que lo modifique, el verbo ser con un sustantivo o un adjetivo. Así, son ejemplos de predicados 'nada', 'está nadando rápidamente', 'es un hombre' 'es pendenciero'. En lógica no se distinguen los predicados de conformidad con las partes gramaticales de que están construidos; de hecho, no hay lugar para verbos, adjetivos y sustantivos comunes. Letras solas representan predicados y no hay símbolos convenidos para adverbios y similares. Además, en la lógica se confiere a los predicados una función más amplia que en el uso común. Por ejemplo, como no se introduce una notación para sustantivos comunes, los nombres comunes que funcionan como sujetos de oraciones se simbolizan por medio de variables y predicados. Considérese, por ejemplo, la oración:

(1) Todo hombre es un animal,

la traducimos así:

(2) Para toda x, si x es un hombre, entonces x es un animal.

En la gramática común, "es un animal" es el predicado de (1). Su traducción (2) tiene el predicado adicional "es un hombre", que

reemplaza al sustantivo común 'hombre' de (1). Usando 'M' para indicar el predicado 'es un hombre' y 'A' para el predicado 'es un animal', podemos entonces representar simbólicamente a (1) como sigue:

(3) Para toda x, $Mx \rightarrow Ax$.

La oración (3) es de hecho un ejemplo de la forma tipo de las oraciones de la especie 'Todo tal y tal es así y así'.

Hay una útil distinción lógica relativa a los predicados, que puede ser ilustrada representando en forma simbólica las oraciones:

(4) Emma era muy alegre

y

(5) El Sr. Knightly era considerablemente de más edad que Emma.

Haciendo

$$e = \text{Emma}$$

y

$$k = \text{Sr. Knightly}$$

y usando '*G*' para el predicado 'era muy alegre' y '*O*' para el predicado 'era considerablemente de más edad que', podemos representar simbólicamente (4) y (5), como sigue:

$$Ge,$$
$$Oke.$$

Por razones obvias llamamos a '*G*' predicado de *una posición* y a '*O*' predicado de *dos posiciones*. Un ejemplo sencillo de predicado de tres posiciones es sugerido por la noción de interposición. Un ejemplo de predicado de cuatro posiciones es sugerido por la relación de dos objetos que distan entre sí lo mismo que otros dos objetos. (Incidentalmente, es posible establecer como axiomas fundamentales de la geometría euclidiana estas dos nociones de interposición y equidistancia).

En la oración:

(6) Emma era muy alegre y bella

el predicado 'era muy alegre y bella' se llama en gramática un predicado *compuesto*. Si simbolizamos con '*B*' el predicado 'era bella', podemos representar (6) como sigue:

(7) $$Ge \;\&\; Be.$$

Por otra parte, si usamos 'W' para 'era alegre y bella', podemos también representar simbólicamente a (6) como sigue.

(8) We.

En lógica no hay compromiso de limitarse a predicados simples y de aquí que la selección entre (7) y (8) se basaría en motivos de comodidad y de contexto. Si en un raciocinio dado la alegría y la belleza están siempre asociadas, entonces (8) sería probablemente el mejor simbolismo. Para un raciocinio que disocie estas dos propiedades, es preferible (7).

§3.4. **Cuantificadores.** Ciertas fórmulas que contienen variables no son verdaderas ni falsas. Así,

(1) x amaba a y

y

(2) $$x + y = z + 2$$

no son verdaderas ni falsas. Sin embargo, si reemplazamos 'x' por 'Emma' y 'y' por 'Sr. Knightley' obtenemos de (1) una oración verdadera (verdadera cuando menos en el mundo de Jane Austen). Y si en (2) reemplazamos 'x' por '2', 'y' por '3' y 'z' por '4', obtenemos una oración falsa sobre aritmética.

Pero uno de los hechos más profundos de la lógica es que no tenemos que reemplazar variables por sustantivos a fin de obtener de (1) y (2) oraciones verdaderas. Otro método igualmente importante consiste en anteponer a expresiones como (1) y (2) frases como 'para toda x', 'existe una x tal que', 'para todas las y'.

Así, de la fórmula

x es un avaro

obtenemos la oración verdadera:

Existe una x tal que x es un avaro,

que tiene el mismo significado que la oración de aspecto menos extraño:

Hay avaros;

y podemos obtener también de esta misma fórmula, la oración falsa:

(3) Para toda x, x es un avaro.

Que tiene el mismo significado que la forma usual.

Todos son avaros.

La frase 'para toda x' recibe el nombre de *cuantificador universal.* El símbolo lógico que usamos para el cuantificador universal, se ilustra escribiendo de nuevo (3) como sigue:

$$(x)\ (x \text{ es un avaro}).$$

En el lenguaje cotidiano, y en matemáticas, palabras y frases como 'alguno', y 'existe una x tal que', 'hay cuando menos una x tal que'. son llamadas *cuantificadores existenciales.* Así pues, la fórmula

$$x > 0$$

Se convierte en una oración verdadera cuando se le antepone el cuantificador existencial 'existe una x tal que'. El símbolo lógico que usamos queda ilustrado formulando la siguiente oración:

$$(\exists x)(x > 0).$$

Es importante advertir que una fórmula que no es verdadera ni falsa, no se vuelve necesariamente verdadera o falsa agregando un solo cuantificador. Debe añadirse un número adecuado de cuantificadores. Por ejemplo, si añadimos a (2) un cuantificador existencial, obtenemos:

(4) $$(\exists x)(x + y = z + 2),$$

y (4) no es verdadera ni falsa. En cambio, si también agregamos '$(\exists y)$' y '$(\exists z)$' entonces obtenemos la oración verdadera:

(5) $$(\exists x)(\exists y)(\exists z)(x + y = z + 2).$$

Puesto en lenguaje matemático común, (5) quiere decir:

Hay números x, y, y z, tales que $x + y = z + 2$.

Como se tiene ahora a mano una notación lógica para los cuantificadores, podemos representar en forma simbólica completa la oración (1) del apartado anterior.

(6) $$(x)(Mx \rightarrow Ax).$$

Como ya se hizo notar, (6) es un ejemplo de una adecuada formulación simbólica de oraciones del tipo 'Todo tal y tal, es así y así' o, lo que es equivalente, oraciones de la especie 'Todos los tales y tales, son así y así'. Así, como un segundo ejemplo, simbolizaríamos la oración

(7) Todos los novatos son inteligentes

como

(8) $$(x)(Fx \rightarrow Ix),$$

donde usamos 'F' para el predicado 'es un novato' e 'I' para el predicado 'es inteligente'.

Considérese ahora la oración:

(9) Ningún novato es inteligente.

Primeramente obtenemos como transcripción parcial de (9):

(10) Para toda x, si x es un novato, entonces x no es inteligente.

y (10) se expresa simbólicamente como:

(11) $$(x)(Fx \rightarrow -Ix).$$

La única diferencia entre (8) y (11) es la presencia en (11) del signo de negación antes del predicado 'I', pero esta diferencia es, por supuesto, definitiva. La oración (11) es un ejemplo de la forma tipo de las oraciones de la especie 'Ningún tal y tal es así y así'.

Veamos ahora algunas oraciones cuyas interpretaciones simbólicas usan cuantificadores existenciales.[1]

(12) Algunos novatos son inteligentes.

Interpretamos (12) como:

(13) Para algunas x, x es un novato y x es inteligente.

Y (13) con facilidad interpreta simbólicamente como:

(14) $$(\exists x)(Fx \,\&\, Ix).$$

Hay un error muy común que cometen los principiantes al interpretar (12), y que debe ser cuidadosamente imitado. La interpretación de (7) y (8) sugiere una falsa interpretación de (12) como:

(15) Para algunas x, si x es un novato, entonces x es inteligente.

La oración 15 se interpreta simbólicamente como:

(16) $$(\exists x)(Fx \rightarrow Ix).$$

Es importante comprender por qué (14) y no (16) es la simbolización correcta de (12). Para cerciorarnos de que (16) es incorrecta, veamos qué es lo que realmente asevera. Si en (16) reemplazamos

[1] Es posible también usar un cuantificador existencial para interpretar (11) como: $-(\exists x)(Fx \,\&\, Ix)$. Mediante el uso de las reglas de inferencia que se dan en el capítulo siguiente, podemos establecer la equivalencia lógica de estas dos interpretaciones.

la fórmula '$Fx \rightarrow Ix$' por la fórmula tautológicamente equivalente* '$-Fx \vee Ix$', obtenemos

(17) $(\exists x)(-Fx \vee Ix)$,

que podemos interpretar verbalmente como

(18) Para algunas x o x no es un novato, o x es inteligente,

o, usando otro giro idiomático para los cuantificadores existenciales, por

(19) Hay algo que es o un no novato, o es inteligente.

Debe resultar obvio que (12) y (19) no son lógicamente equivalentes. Si todos los novatos fueran completamente estúpidos, (12) sería falsa, mientras que (19) sería todavía verdadera, mientras hubiera algo en el universo que no fuera un novato.

Las aseveraciones de la misma forma de (16) son casi siempre verdaderas, pero completamente triviales. Consecuentemente, muy raras veces son una interpretación correcta de cualquier oración del lenguaje ordinario en que estemos interesados.

Otro ejemplo puede ayudar a subrayar la trivialidad de (16). Interpretemos la oración

(20) Algunos hombres tienen tres cabezas.

Su correcta representación simbólica es.

(21) $(\exists x)(Mx \;\&\; Tx)$,

y (20) y (21) son, que nosotros sepamos, falsas. Si hubiéramos interpretado (15) como

$$(\exists x)(Mx \rightarrow Tx),$$

y con esto hubiéramos optado por el enunciado lógicamente equivalente:

(22) $(\exists x)(-Mx \vee Tx)$,

---

* La equivalencia tautológica puede demostrarse con la siguiente tabla de verdad abreviada:

(1) $Fx \rightarrow Ix$	(2) $Fx \vee Ix$	(3) (1) $\leftrightarrow$ (2)
T	T	T
F	F	T
T	T	T
T	T	T

En los dos primeros casos *Fx* es verdadera y en los dos segundos es falsa. Se supone que el lector tiene ya suficiente experiencia para suplir los detalles faltantes. (N. del T.)

estaríamos afirmando el absurdo de que (20) es verdadero, mientras haya en el universo objetos que no sean hombres, pues esta trivial condición garantiza la verdad de (22).

De la discusión anterior debe resultar la evidencia de la correcta interpretación de oraciones relacionadas como

Algunos novatos no son inteligentes. Esa correcta interpretación es

$$(\exists x)(Fx \;\&\; -Ix).$$

Una útil ampliación de las clases de oraciones estudiadas hasta ahora, es a oraciones que usan variables y nombres propios. Así, representamos simbólicamente

Adams no está casado con ninguna,

por

$$(x)\text{-}(Max),$$

donde '*Max*' se lee Adams está casado con *x*. Como otro ejemplo, representamos simbólicamente:

Algunos novatos salen a pasear con Elizabet

por

$$(\exists x)(Fx \;\&\; Dxe),$$

donde el predicado de una posición '*F*' y el de dos posiciones '*D*' tinene un significado obvio y e = Elizabeth.

EJERCICIOS

1. Exprese sin usar variables las siguientes dos oraciones verdaderas de aritmética:

(a) $(\exists x)(\exists y)(x + y \neq x \cdot y)$.
(b) $(\exists x)(\exists y)(\exists z)(x - (y - z) \neq (x - y) - z)$.

2. Transcriba las siguientes oraciones del lenguaje cotidiano, a una notación que use términos, predicados, cuantificadores y conectivos oracionales. Use para los predicados y los nombres propios las letras que se indican. Nótese que los sustantivos plurales no precedidos de 'algunos' o 'todos', algunas veces deben ser transcritos usando cuantificadores universales y otras veces cuantificadores existenciales.

(a) Todos los alumnos de último año son dignos. (Sx, Dx)
(b) Algunas alumnas de penúltimo año son bonitas. (Jx, Px)
(c) Ningún novato es serio. (Fx)
(d) Algunas novatas no son bonitas.
(e) Algunas alumnas de último año son bonitas y serias.
(f) Betty es bonita, pero no es seria. (b)
(g) No todas las novatas son bonitas.
(h) Algunos alumnos de último año que no son serios, gustan del griego. (Tx, Gx)

(i) Ningún novato no es serio.
(j) Algunas alumnas de penúltimo año que gustan del griego no son bonitas.
(k) Todas las novatas que gustan del griego son bonitas.
(l) Cualquiera que es aficionado al griego es o de último o de penúltimo año.
(m) Algunos alumnos de último año son aficionados al griego y a las matemáticas. (Mx).
(n) Ningún alumno de penúltimo año es aficionado al griego y a las matemáticas simultáneamente.
(o) Hay algunos alumnos de último año que son aficionados al griego pero no a las matemáticas.
(p) Todo novato sale a pasear con alguna alumna de penúltimo año. (Dxy).
(q) Elizabeth no sale a pasear con ningún novato. (e).
(r) Algunos alumnos de penúltimo año sólo salen a pasear con alumnas de último año.
(s) Algunos novatos salen a pasear con alumnas de último año.
(t) Algunos alumnos de último año no salen a pasear con alumnas de penúltimo.
(u) Tanto alumnos de último como de penúltimo año salen a pasear con Betty.
(v) Algunos novatos salen a pasear sólo con alumnas de último año aficionadas al griego.
(w) Solamente novatos salen a pasear con novatas.
(x) Los novatos salen a pasear sólo con novatas.
(y) Adams sale a pasear con novatas solamente si son bonitas. (a).
(z) Adams sale a pasear con una alumna de penúltimo año.

§3.5. **Variables ligadas y libres.** La incidencia de una variable en una fórmula puede ser de uno de dos tipos: o está esa incidencia regulada por un cuantificador, o no lo está. Para el enunciado exacto de las reglas de inferencia aplicables a los cuantificadores es importante precisar esta noción de que una variable "está regulada" por un cuantificador. Para comenzar, necesitamos hacer más precisa la noción de fórmula. Una *fórmula atómica* es un predicado seguido o flanqueado por el número apropiado de términos como argumentos. Así, las expresiones

$$Ra, \qquad x + y = z, \qquad x \text{ es azul}$$

son fórmulas atómicas. Podemos dar entonces lo que se llama una definición *recursiva de fórmulas*:

(a) *Toda fórmula atómica es una fórmula.*
(b) *Si* **S** *es una fórmula, entonces* $-(\mathbf{S})$ *es una fórmula.*
(c) *Si* **R** *y* **S** *son fórmulas, entonces* $(\mathbf{R}) \,\&\, (\mathbf{S})$, $(\mathbf{R}) \vee (\mathbf{S})$, $(\mathbf{R}) \rightarrow (\mathbf{S})$, *y* $(\mathbf{R}) \leftrightarrow (\mathbf{S})$ *son fórmulas.*
(d) *Si* **R** *es una fórmula y* **v** *es cualquier variable, entonces* $(\mathbf{v})(\mathbf{R})$ *y* $(\exists\, \mathbf{v})(\mathbf{R})$ *son fórmulas.*

(e) *Ninguna expresión es una fórmula a menos que el que lo sea se siga de las reglas anteriores.*

Esta definición hace explícito lo que debería ser intuitivamente obvio.[2] (En la práctica omitimos los paréntesis indicados cuando son innecesarios). Dar esta definición equivale a incluir entre las reglas de ajedrez, damas o bridge, la indicación de cuáles son las posiciones que tienen sentido en el juego; por ejemplo, ningún jugador puede tener seis cartas y su compañero ocho, en ningún momento del juego de bridge. Y la expresión

$$x = y \rightarrow \vee(x) \leftrightarrow$$

no es una fórmula, sino una cadena de símbolos, que carece de significado, esto es, que no es significativa. (Aquí la expresión no significativa debe tomarse en sentido formal. No se requiere que una fórmula sea significativa en el sentido de que diga algo significativo, simplemente tiene que tener cierta estructura formal o *sintáctica.*)

Ya en posesión de una idea relativamente precisa de fórmula, podemos ahora definir la importante noción de *alcance* de un cuantificador.

DEFINICIÓN. *El* ALCANCE* *de un cuantificador que aparece en una fórmula, es el cuantificador, juntamente con la fórmula mínima que sigue al cuantificador.*

En los ejemplos siguientes el alcance del cuantificador '$(\exists x)$' está indicado por la línea trazada debajo de la fórmula.

(1) $$(\exists x)Mx \vee Rx$$

(2) $$(y)(\exists x)(x > y \,\&\, (z)(z = 2))$$

alcance

(3) $$(\exists x)(y)((xy = 0) \,\&\, (x)(z)(x + z = z + x))$$

alcance

(4) $$(\exists x)(y)(xy = 0) \,\&\, (x)(z)(x + z = z + x)$$

alcance

[2] La definición que se da aquí no es completamente precisa, pues no se da una lista completa de términos y predicados primitivos, como se necesitaría para construir un lenguaje completamente formalizado. Además, en un lenguaje formalizado que admite términos como "$x + y$", se necesitaría también una definición recursiva de los términos.

* Consecuentes con el propósito manifestado de mantenernos fieles al pensamiento del autor y a sus modos y medios de expresión, hemos traducido "scope" por *alcance* y no por *extensión*, en virtud de que este segundo término hace referencia a un concepto perfectamente definido de la lógica aristotélica, con el cual no coincide el que aquí se define. Tampoco podemos usar el término "dominio", porque el autor lo usa con un significado más específico. (Véase Pág. 98). (N. del T.)

El paréntesis en una fórmula siempre aclara en una forma perfectamente natural qué expresión es la fórmula mínima que sigue a un cuantificador. Así, (1) 'Mx' es la fórmula mínima que sigue a '$(\exists x)$'. En (2), el cierre del paréntesis que se abre inmediatamente después de '$(\exists x)$' ocurre al extremo de la fórmula, lo que determina que la fórmula mínima sea en este caso todo el resto de la fórmula. Pero como lo muestra (3), no basta con buscar el cierre del paréntesis que se abre inmediatamente después del cuantificador, pues en el caso de (3) esto determinaría falsamente a '(y)' como la fórmula mínima que sigue a '$(\exists x)$', pero el cuantificador '(y)' no es una fórmula. Sin embargo, si se pasan por alto esos cuantificadores, siempre es eficaz la regla de buscar el cierre del paréntesis.

Ahora estamos en condiciones de definir las incidencias *ligadas* y *libres* de las variables.

DEFINICIÓN. *La incidencia de una variable en una fórmula es* LIGADA *si y sólo si esta incidencia queda dentro del alcance de un cuantificador que usa esta variable.*

El significado de la frase 'cuantificador que usa esta variable' debe quedar claro. Así, el cuantificador '$(\exists x)$' usa la variable 'x'; el cuantificador '(y)' usa la variable 'y'. En las fórmulas (2) a (4) anteriores, todas las incidencias de las variables son ligadas.

DEFINICIÓN. *Una incidencia de una variable es* LIBRE *si y sólo si esta incidencia de la variable no está ligada.*

En la fórmula:

(5) $$(\exists y)(x > y)$$

ambas incidencias de la variable 'y' son ligadas y la incidencia única de 'x' es libre. En (1) las primeras dos incidencias de 'x' son ligadas y la tercera incidencia es libre.

DEFINICIÓN. *Una variable es una* VARIABLE LIBRE *en una fórmula si y sólo si cuando menos una incidencia de la variable es libre, y una variable es una* VARIABLE LIGADA *en una fórmula si y sólo si cuando menos una incidencia de la variable es ligada.*

Notamos que una variable puede ser tanto libre como ligada en la misma fórmula, pero cualquier incidencia dada de la variable es ligada o libre, pero no ambas cosas. Así, en la fórmula

$$y > 0 \vee (\exists y)(y < 0)$$

la variable 'y' es tanto libre como ligada; su primera incidencia es libre y la segunda y la tercera ligadas.

Sin advertirlo explícitamente, hemos reservado la palabra 'oración' para aplicarla a fórmulas que son verdaderas o falsas, pero una fórmula es verdaera o falsa si y sólo si no tiene variables libres. Consecuentemente, una precisa definición sintáctica es simplemente:

DEFINICIÓN... *Una* ORACION *es una fórmula que no tiene variables libres.*

Así, las fórmulas (2) a (4) de este aparato son oraciones, pero (5) no lo es, a causa de la ocurrencia libre de 'x'.

Con respecto al problema de simbolizar correctamente oraciones del lenguaje cotidiano, debe hacerse notar que el resultado final no debe contener variables libres. En otras palabras, la expresión simbolizada debería ser también una oración. En el Cap. 7, por otra parte, discutiremos la tendencia universal de usar fórmulas que no tiene variables libres.

EJERCICIOS

1. Clasifíquense las siguientes expresiones como: (i) términos, (ii) fórmulas, o (iii) ni términos ni fórmulas.

(a) Thomas Jefferson
(b) lentamente
(c) el número seis
(d) x es púrpura
(e) el entero y
(f) el número x, tal que $x = x + x$
(g) $x + y = z$
(h) $(x + y) + z^2$
(i) todo hombre x
(j) ese hombre que será el próximo presidente
(k) x odia a y, pero ama a z
(l) $(3 + 1) + 10$
(m) muy bonito

2. ¿Cuántas incidencias libres de variables y cuántas variables libres hay en cada una de las fórmulas siguientes? ¿Cuál de las fórmulas son oraciones?

(a) $(x)(y)(z)(x > y \; \& \; y > z) \rightarrow (\exists w)(w > w)$
(b) $(\exists x)(x$ es roja$) \vee (y)(y$ es azul y x es púrpura$)$
(c) $x + x = x + x$
(d) $(\exists y)(x + x = x + x)$
(e) $(\exists x)(\exists y)(x$ está casada con y y z es su hijo$)$

§3.6. Un ejemplo final. Las oraciones consideradas hasta aquí en este capítulo son de estructura relativamente sencilla. Es deseable considerar un ejemplo relativamente más complicado:

(1) Si un instante ocurre después de un segundo instante, entonces hay un instante después del segundo y antes del primero.

El mejor modo de comenzar una transcripción es el de introducir variables antes de representar simbólicamente la oración. Está claro que necesitamos tres variables distintas al transcribir (1):

(2) Si el instante x ocurre después del instante y, entonces hay un instante z, tal que z está después de y y z está antes de x.

No es difícil pasar de (2) a una completa representación simbólica de (1). El cambio principal consiste en reemplazar 'instante' por el predicado correspondiente. El resultado de esta sustitución es una oración más burda, pero cuya estructura lógica es más sencilla que la de (2).

(3) Si x es un instante, y es un instante, y x está después de y, entonces existe una z tal que z es un instante, z está después de y y z está antes de x.

Resulta ahora un trabajo rutinario representar simbólicamente a (3):

(4) $Tx \mathbin{\&} Ty \mathbin{\&} Axy \rightarrow (\exists z)(Tz \mathbin{\&} Azy \mathbin{\&} Bzx)$,

donde 'T' simboliza 'es un instante', 'A' el predicado de dos posiciones 'esta' después de y 'B' el predicado de dos posiciones 'está antes de'.

Proceden algunas observaciones con respecto de las fórmulas (1) a (4). La fórmula (4) tiene dos variables libres y, en consecuencia, no es estrictamente una oración que sea verdadera o falsa. Por otra parte, parece natural e intuitivamente correcto considerar a (1) como una oración. En (1) las palabras 'un' y 'segundo' funcionan como cuantificadores universales. Se tendrá quizás una mejor percepción de su función si se inserta la palabra 'cualquiera' después de 'un instante' y de 'un segundo instante' y se deja la redacción de (1) como sigue:

> Si un instante cualquiera ocurre después de un segundo instante cualquiera, entonces. . .

Por otra parte, obtenemos fácilmente una oración (en el estricto sentido de la palabra) de (4) añadiendo dos cuantificadores universales:

(5) $(x)(y)[Tx \& Ty \& Axy \rightarrow (\exists z)(Tz \& Azy \& Bzx)]$

que es el paso final en la completa representación simbólica de (1).

La comparación de (1) con (5) pone de relieve el hecho de que los cuantificadores universales se expresan en el lenguaje usual por medio de diversos modismos. Podría parecer difícil seguir la pista de estos modismos e identificarlos, pero en la práctica resulta bastante sencillo: después de obtener una transcripción que parezca intuitivamente adecuada, agréguense tantos cuantificadores universales como sean necesarios para obtener una oración, como hicimos al derivar (5) de (4). Como ya se hizo notar, en el capítulo siguiente veremos que fórmulas como (4), que contienen variables libres y que no son oraciones, desempeñan un importante papel en la teoría de la inferencia.

Quizá sea necesario repetir que no es posible dar reglas precisas para representar simbólicamente las oraciones del lenguaje usual. La corrección o incorrección de una transcripción propuesta debe decidirse por medio de diversas consideraciones intuitivas e informales. En ciertos casos, parecen igualmente correctas varias transcripciones no equivalentes, pero generalmente sucede que si se producen dos transcripciones no equivalentes, puede demostrarse que una es incorrecta.

EJERCICIOS

1. Usando las letras indicadas para los predicados y las constantes, transcríbase lo siguiente. Las variables que siguen a los predicados no son necesariamente las que se usarían en la transcripción simbólica de la oración.

(a) Los alumnos de segundo año con aficionados al griego sólo si son aficionados a las matemáticas. (Sx, Gx, Mx)
(b) Los alumnos de último año sólo salen a pasear con alumnas del penúltimo. (Sx, Dxy, Jx)
(c) Algunos alumnos de último año son aficionados al griego, pero ningún alumno de último año es aficionado simultáneamente al francés y a las matemáticas. (Fx)
(d) Todo instante ocurre después de algún instante. (Ix, Axy)
(e) Si dos instantes no son idénticos, entonces uno ocurre después del otro. ($x = y$).
(f) No hay un instante tal que todo instante ocurra después de él.
(g) No hay un instante tal que no ocurra después de él ningún instante.
(h) Si un instante ocurre después de un segundo instante, entonces el segundo ocurre antes del primero. (Bxy)
(i) Los únicos alumnos de segundo año que salen a pasear con Betty son los aficionados al griego. (Dxy, b)
(j) Algunos alumnos de último año que son aficionados a las matemáticas no salen a pasear con Elizabeth. (e)

(k) Algunos alumnos de primer año que no son aficionados al francés, salen a pasear con Betty y también con Elizabeth.

(l) Algunos alumnos de primer año que son aficionados al griego y a las matemáticas no salen a pasear con Betty ni con Elizabeth.

(m) Si a todos los alumnos de segundo año les gusta el griego, entonces a algunos de primero también.

(n) O a todos los alumnos de primer año les gustan las matemáticas, o algunos de segundo son aficionados al griego.

(o) Si alumnos de último año salen a pasear solamente con alumnas de primero, entonces algunos alumnos de último año no salen a pasear con nadie.

(p) Si ningun alumno de último año sale a pasear con Betty, entonces sale alguno de segundo o alguno de penúltimo.

2. Usando la notación lógica desarrollada y los símbolos normales de aritmética, como '$<$' para 'menor que', hágase la transcripción simbólica de las siguientes oraciones.

(a) Hay un número x menor que 5 y mayor que 3.

(b) Dado cualquier número x, hay un número menor y.

(c) No existe un número máximo.

(d) Para dos números cualesquiera x y y, la suma de x y de y es igual a la suma de y y de x.

(e) Existen números x, y y z, tales que la diferencia de x y de y es menor que el producto de x y de z.

(f) Para todo número x hay un número y tal que para todo número z, si la diferencia de z y de 5 es menor que y, entonces la diferencia de x y de 7 es menor que 3.

(g) Si es igual a cero la suma de dos números, ninguno de los cuales es igual a cero, entonces uno de los números es mayor que cero.

CAPITULO 4

TEORIA GENERAL DE LA INFERENCIA

§4.1. La inferencia, cuando implica sólo cuantificadores universales. Es conveniente aprender poco a poco las reglas un tanto complejas de la inferencia, que van más allá de los métodos alético-funcionales del Cap. 2 y se apoyan en las propiedades lógicas de los cuantificadores. Se necesitan cuatro nuevas reglas: un par para suprimir y agregar cuantificadores universales y otro par semejante para cuantificadores existenciales. En términos generales, la *estrategia* para manejar inferencias que impliquen cuantificadores, se desarrolla en cuatro partes:

I. Representación simbólica de las premisas en notación lógica.
II. Supresión de cuantificadores, de conformidad con las reglas introducidas en este capítulo.
III. Aplíquense métodos oracionales de deducción, para obtener conclusiones sin cuantificadores.
IV. Añádanse cuantificadores para obtener una conclusión en forma final.

(Este enunciado sobre estrategia debe tomarse como una sugestión útil; algunas veces no es aplicable).

A fin de comenzar en forma sencilla, en este apartado estudiaremos fórmulas que contienen sólo cuantificadores universales con alcances que abarcan toda la fórmula. Un silogismo clásico puede servir como primer ejemplo.

Todos los animales son mortales.
Todos los seres humanos son animales.
Luego todos los seres humanos son mortales.

Usando los métodos de transcripción del capítulo anterior, represen-

tamos simbólicamente este raciocinio como sigue:

$$\begin{array}{l} (x)(Ax \rightarrow Mx) \\ (x)(Hx \rightarrow Ax) \\ \hline (x)(Hx \rightarrow Mx) \end{array}$$

Para construir una deducción correspondiente a este sencillo raciocinio, introducimos dos nuevas reglas de inferencia. La primera nos permite suprimir un cuantificador universal, cuyo alcance se extiende a toda la fórmula. La idea intuitiva en que se basa la regla es que cualquier cosa que sea cierta para todo objeto, es cierta para cualquier objeto dado. La segunda regla nos permite agregar un cuantificador; en este caso inferimos una verdad aplicable a todos los objetos, de una verdad aplicable a un objeto elegido arbitrariamente.

La deducción se desarrolla como sigue:

EJEMPLO 1.

{1}	(1) $(x)(Ax \rightarrow Mx)$	P
{2}	(2) $(x)(Hx \rightarrow Ax)$	P
{1}	(3) $Ax \rightarrow Mx$	Supresión del cuantificador universal de (1).
{2}	(4) $Hx \rightarrow Ax$	Supresión del cuantificador universal de (2).
{1, 2}	(5) $Hx \rightarrow Mx$	3, 4 T (Ley del Silogismo Hipotético).
{1, 2}	(6) $(x)(Hx \rightarrow Mx)$	Agréguese el cuantificador universal a (5).

El análisis de este ejemplo en función de la estrategia de cuatro tiempos sugerida anteriormente, es evidente:

PASO I. *Simbolización de las premisas*: líneas (1) y (2).
PASO II. *Supresión de cuantificadores*: líneas (3) y (4).
PASO III. *Aplicación de métodos oracionales*: línea (5).
PASO IV. *Adición del cuantificador*: línea (6).

Deseamos ahora enunciar exactamente las dos reglas que rigen los cuantificadores universales. Llamamos a la primera la regla de la *especificación universal*, a causa de que un enunciado verdadero sobre

toda cosa es verdadero de cualquier cosa arbitrariamente elegida o *especificada*.[1] Su formulación exacta es:

> *Regla de la Especificación Universal*: US. *Si una fórmula* **S** *resulta de una fórmula* **R**, *sustituyendo por un término* **t** *cada incidente libre de una variable* **v** *en* **R**, *entonces* **S** *es deducible de* (**v**)**R**.[2]

En la deducción anterior, el término **t** es simplemente 'x' en ambos casos de especificación universal y por eso hablamos de *suprimir* cuantificadores, pero en general, deseamos estar en aptitud de reemplazar la variable cuantificada por cualquier término.[3] Así, si 'a' es el nombre de John Quincy Adams, podemos inferir '$Aa \rightarrow Ma$', así como '$Ay \rightarrow My$', de '$(x)\ (Ax \rightarrow Mx)$'. Y en el caso de aritmética, del enunciado '$(x)\ (x + 0\, x)$' podemos inferir por especificación universal '$(x + y) + 0 = x + y$'; en este caso el término **t** es '$x + y$'.

La segunda regla, que permite la adición de cuantificadores, se complica por la necesidad de una restricción de su ámbito de aplicabilidad, aun para los limitados argumentos de este apartado. El carácter de esta restricción puede dilucidarse por el siguiente sencillo argumento. Comenzamos con la premisa de que una x cualquiera es humana y concluimos falazmente que todo es humano.

{1}	(1) Hx	P
{1}	(2) $(x)Hx$	Agréguese a (1) el cuantificador universal.

El raciocinio dice que para todo objeto x, si x es humano, entonces todo es humano. Si ningún objeto fuera humano, el antecedente seria siempre falso, pero puesto que hay humanos, hay casos de x para los cuales el antecedente es verdadero y, obviamente, la conclusión es falsa. Pero no es difícil de encontrar una restricción que cierre el camino a esta falaz inferencia: no se *generalice universalmente*, esto es, no se cuantifique universalmente una variable que está libre en una

[1] Se usa a menudo la expresión "instanciación o ejemplificación universal". Yo prefiero "especificación" en vez de "instanciación", sobre la base de que sólo como último recurso deben introducirse tecnicismos no usados en el lenguaje cotidiano.

[2] Aquí y en las reglas subsecuentes, a fin de evitar cualquier posible ambigüedad en el alcance del cuantificador, podemos tratar 'R' como si estuviera flaqueada por paréntesis. Por ejemplo, leemos 'S es deducible de (v)R', como si fuera 'S es deducible de (v) (R)'.

[3] En lo que respecta a la consideración sistemática de los términos en la teoría de la deducción, hay una restricción informal que es necesario aclarar. Siempre suponemos que el objeto designado por un término existe y es único. En otras palabras, se excluyen los nombres y las descripciones de objetos ficticios. Se prohiben, por tanto, términos como 'el hijo único de Franklin Delano Roosevelt', 'la undécima esposa de Napoleón', 'el número primo mayor que ocho y menor que diez'. En el Cap. 8 trataremos de nuevo estos temas.

premisa. En el anterior argumento falaz generalizamos universalmente la variable '*x*', que está libre en la premisa '*Hx*'. Para fijar las ideas, diremos que una variable libre de una premisa es *imprecisa*. Además, una variable que es libre en una premisa es también *imprecisa* en cualquier línea en que es libre y que depende de la premisa. Listamos las variables imprecisas en una fórmula, a la derecha, en una deducción (véase el Ej. 2 posteriormente).

El enunciado exacto de la regla de inferencia para añadir un cuantificador universal es entonces:

> *Regla de Generalización Universal*: *UG*. *De la fórmula* **S** *podemos deducir* (**v**)**S**, *siempre que la variable* **v** *no sea débil en* **S**.[4]

El siguiente ejemplo, que se aplica a las dos nuevas reglas de inferencia, ilustra tres cosas: (i) la notación abreviada: *US* para la especificación universal [líneas (4) a (6)] y la notación abreviada: *UG* para la generalización universal [línea (12)]; (ii) cómo se indican las variables débiles [líneas (7) a (10)]; (iii) cómo se aplica la especificación universal cuando acontece que las premisas están simbolizadas mediante el uso de diferentes variables.

Ejemplo 2 (Lewis Carrol). *Ningún pato quiere bailar. No hay ningún oficial que no quiera bailar. Todas mis aves de corral son patos. En consecuencia, ninguna de mis aves de corral son oficiales.*

{1}	(1) $(x)(Dx \rightarrow -Wx)$	P
{2}	(2) $(y)(Oy \rightarrow Wy)$	P
{3}	(3) $(z)(Pz \rightarrow Dz)$	P
{1}	(4) $Dx \rightarrow -Wx$	1 *US*
{2}	(5) $Ox \rightarrow Wx$	2 *US*
{3}	(6) $Px \rightarrow Dx$	3 *US*
{7}	(7) Px	*x* P
{3, 7}	(8) Dx	*x* 6, 7 T
{1, 3, 7}	(9) $-Wx$	*x* 4, 8 T
{1, 2, 3, 7}	(10) $-Ox$	*x* 5, 9 T
{1, 2, 3}	(11) $Px \rightarrow -Ox$	7, 10 C.P.
{1, 2, 3}	(12) $(x)(Px \rightarrow -Ox)$	11 *UG*

Nótese que '*x*' es imprecisa por primera vez en la línea (7), porque es libre en '*Px*' y '*Px*' es una premisa. Sigue imprecisa en las líneas (8) a (10), porque estas tres líneas dependen de (7), como lo indican los números listados a la izquierda. La condicionalización resultante en

[4] La restricción por imprecisión se discute también en el apartado siguiente. Su significado intuitivo se expone ulteriormente en el Cap. 7.

la línea (11) termina la imprecisión de 'x' y permite la generalización universal de (11). Al elaborar deducciones, es un error preocuparse demasiado por la restricción por imprecisión. Generalmente, se está dispuesto a generalizar una variable cuando hacerlo es al mismo tiempo natural y correcto. De aquí que la recomendación práctica sea: constrúyase una deducción sin tomar en cuenta la imprecisión, al terminar, inspecciónese rápidamente cada línea, márquese la imprecisión de las variables que corresponda y véase que ninguna variable débil haya sido generalizada universalmente. En las deducciones que impliquen manipulación de cuantificadores casi cualquier aplicación de la regla de la prueba condicional implica marcado rutinario de imprecisión de cuando menos una variable, para unas cuantas líneas.

La representación simbólica de las tres premisas del argumento anterior, mediante el uso de tres variables diferentes, se hizo sólo con fines de ilustración y muestra cómo el término 'x' es sustituido en las tres por especificación universal. Comúnmente, lo que hay que hacer es usar la misma variable tanto como sea posible; en este caso usar 'x' para simbolizar las tres premisas.

Un análisis de la estrategia usada en el Ej. 2 muestra el mismo desarrollo en cuatro pasos del Ej. 1.

PASO I. *Simbolización de las premisas*: líneas (1) a (3).
PASO II. *Supresión de cuantificadores*: líneas (4) a (6).
PASO III. *Aplicación de métodos oracioncles*: líneas (7) a (11).
PASO IV. *Agréguese el cuantificador*: línea (12).

Debe resultar obvio que, dentro de la restricción de este apartado (cuando se tienen al frente sólo cuantificadores universales), el uso de la sagacidad lógica se concentra casi totalmente en el Paso III, aplicación de métodos oracionales. Nunca se insistirá demasiado en la importancia general de un firme dominio de las implicaciones tautológicas.

Un ejemplo final en este apartado será útil para ilustrar las inferencias que contienen nombres propios.

EJEMPLO 3. *Ningún episcopaliano o presbiteriano es unitario. John Quincy Adams era unitario. En consecuencia, no era episcopaliano.*

{1}	(1) $(x)[(Ex \vee Px) \rightarrow -Ux]$	P
{2}	(2) Ua	P
{1}	(3) $(Ea \vee Pa) \rightarrow -Ua$	1 *US*
{1, 2}	(4) $-(Ea \vee Pa)$	2, 3 T
{1, 2}	(5) $-Ea \;\&\; -Pa$	4 T
{1, 2}	(6) $-Ea$	5 T

Deducimos la línea (3) de la (1) por *US*, reemplazando '*x*' por el nombre propio '*a*', donde a = John Quincy Adams. Nótese la aplicación de la ley de De Morgan para obtener (5) de (4).

EJERCICIOS

Constrúyase, de ser posible, una deducción que corresponda a los argumentos siguientes. Las variables usadas para indicar el número de posiciones de los predicados no son necesariamente las variables que han de usarse para simbolizar las premisas. Para la representación simbólica de todas las oraciones sólo se requieren cuantificadores universales colocados al frente.

1. Todos los científicos son racionalistas. Ningún filósofo británico es racionalista. En consecuencia, ningún filósofo británico es científico. (Sx, Rx, Bx)

2. Ningún existencialista aprecia a los positivistas. Todos los miembros del Círculo de Viena son positivistas. En consecuencia, ningún existencialista aprecia a ningún miembro del Círculo de Viena. (Ex, Lxy, Px, Mx)

3. Si una primera persona es padre de una segunda, entonces la segunda no es padre de la primera. En consecuencia, ningún hombre es su propio padre. (Fxy)

4. Para toda x y y, o x es cuando menos tan pesada como y, o y es cuando menos tan pesada como x. En consecuencia, x es cuando menos tan pesada como sí misma. (Hxy)

5. Adams es un muchacho que no tiene carro. María sale a pasear sólo con muchachos que tienen carro. En consecuencia, María no sale a pasear con Adams. (Bx, Ox, Dxy, a, m)

6. Todo miembro del Ayuntamiento vive dentro de los límites de la ciudad. El señor Fairman no vive dentro de los límites de la ciudad. En consecuencia, el señor Fairman no es un miembro del Ayuntamiento. (Mx, Lx, f)

7. Todo miembro del comité político es demócrata o republicano. Todo miembro demócrata del comité político vive en California. Todo miembro republicano del comité político es miembro del comité de impuestos. En consecuencia, todo miembro del comité político que no forma parte también del comité de impuestos, vive en California. (Px, Dx, Rx, Lx, Tx,)

8. Cualquiera que trabaje en la fábrica, o está sindicalizado, o tiene un puesto de confianza. Adams no está sindicalizado y tampoco tiene un puesto de confianza. En consecuencia, Adams no trabaja en la fábrica. (Wx, Ux Mx, a)

9. Ptah es un dios egipcio y es padre de todos los dioses egipcios. En consecuencia, es padre de sí mismo. (p, Gx, Fxy)

10. Si se tiene: (i) para cualesquiera número x, y y z, si $x > y$ y $y > z$, entonces $x > z$; (ii) para cualquier número x, no sucede que $x > x$. En consecuencia, para dos números cualesquiera x y y, si $x > y$, entonces no sucede que $y > x$. (Nx, $x > y$)

11. En la teoría del comportamiento racional, que tiene aplicaciones en economía, ética y psicología, es de importancia al concepto de individuo que prefiere un objeto o estado de cosas a otro. Podemos decir que un individuo prefiere *débilmente* x a y, si no prefiere estrictamente y a x. Usamos el

concepto de preferencia débil por comodidad formal, pues si usamos preferencias estrictas necesitaríamos también un concepto de *indiferencia*. La medula del presente ejercicio consiste en la demanda de que, sobre la base de las dos sencillas propiedades postuladas de preferencia débil y las definiciones adecuadas de preferencia estricta e indiferencia, en función de preferencia débil, se demuestra que es posible inferir lógicamente todas las propiedades esperadas de preferencia estricta y de indiferencia. Usese 'Q' para preferencia débil, 'P' para preferencia estricta e 'I' para indiferencia. Los dos postulados o premisas sobre el predicado 'Q' dicen solamente que es transitivo, y que de dos objetos cualesquiera del dominio de objetos que se consideran, uno es preferido débilmente al otro. En notación simbólica:

(1) $(x)(y)(z)(xQy \,\&\, yQz \rightarrow xQz)$,
(2) $(x)(y)(xQy \vee yQx)$.

Como premisas adicionales introduciremos las dos definiciones obvias:

(3) $(x)(y)(xIy \leftrightarrow xQy \,\&\, yQx)$,
(4) $(x)(y)(xPy \leftrightarrow -yQx)$.

Dedúzcanse las conclusiones siguientes de estas cuatro premisas:

(a) $(x)(xIx)$
(b) $(x)(y)(xIy \rightarrow yIx)$
(c) $(x)(y)(z)(xIy \,\&\, yIz \rightarrow xIz)$
(d) $(x)(y)(xPy \rightarrow -yPx)$
(e) $(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$
(f) $(x)(y)(xIy \rightarrow -(xPy \vee yPx))$
(g) $(x)(y)(z)(xIy \,\&\, yPz \rightarrow xPz)$
(h) $(x)(y)(z)(xIy \,\&\, zPx \rightarrow zPy)$

12. E. V. Huntington dio una lista de axiomas para la parte informal de la obra *Principia Mathematica* de Whitehead y Russell, que descansa en tres símbolos primitivos, un símbolo de una sola posición 'C' de relación, un símbolo de operación binaria '$+$', y un símbolo de operación unitaria '$^{\text{I}}$'.[5] La principal interpretación que se intenta de la teoría es que las variables toman como valores proposiciones elementales (u oraciones, si se desea); una proposición tiene la propiedad C si es verdadera, $x + y$ es la disyunción de las dos proposiciones x y y; x^{I} es la negación de la proposición x.[6] (Así pues, este conjunto de postulados está íntimamente conectado con la lógica de los conectivos oracionales tratada en el Cap. 1.)

Los cinco axiomas son:

A1. $(x)(y)[C(x + y) \rightarrow C(y + x)]$
A2. $(x)(y)[C(x) \rightarrow C(x + y)]$
A3. $(x)[C(x^{\text{I}}) \rightarrow -C(x)]$
A4. $(x)[-C(x^{\text{I}}) \rightarrow C(x)]$
A5. $(x)(y)[C(x + y) \,\&\, C(x^{\text{I}}) \rightarrow C(y)]$

Constrúyanse deducciones para demostrar que los teoremas siguientes son consecuencias logicas de las axiomas. Puesto que algunos de los teoremas son

[5] *Boletín de la American Mathematical Society*, Vol. 40 (1934), Págs. 127-136. Nuestra formalización reduce el número de axiomas de ocho a cinco.

[6] Aquí el símbolo '+' no tiene ninguna conexión con la aritmética común.

útiles para demostrar otros, se recomienda una deducción larga para todas las partes. (Posteriormente estableceremos reglas para insertar resultados demostrados anteriormente, en nuevas deducciones.) Si se da una deducción larga, indíquese al final qué línea de la deducción corresponde a cada teorema. Cerciórese de que ningún teorema depende de más de los cinco axiomas y premisas dadas.

Teorema 1.	$(x)[-C(x) \rightarrow C(x^{1})]$
Teorema 2.	$(x)[C(x) \rightarrow -C(x^{1})]$
Teorema 3.	$(x)(y)[-C(x+y) \rightarrow -C(y+x)]$
Teorema 4.	$(x)(y)[C(x+y) \mathbin{\&} -C(x) \rightarrow C(y)]$
Teorema 5.	$(x)(y)[-C(x+y) \rightarrow -C(x) \mathbin{\&} -C(y)]$
Teorema 6.	$(x)(y)[-C(x) \mathbin{\&} -C(y) \rightarrow -C(x+y)]$
Teorema 7.	$(x)[C((x^{1})^{1}) \leftrightarrow C(x)]$
Teorema 8.	$(x)(y)[C(x) \mathbin{\&} C(x^{1}+y) \rightarrow C(y)]$
Teorema 9.	$(x)C((x+x)^{1}+x)$
Teorema 10.	$(x)(y)C(y^{1}+(x+y))$
Teorema 11.	$(x)(y)C((x+y)^{1}+(y+x))$
Teorema 12.	$(x)(y)(z)C[(x^{1}+y)^{1}+[(z+x)^{1}+(z+y)]]$ [7]

§ 4.2. **Interpretaciones y validez.** Al principio del Cap. 2, se introdujo la noción de *interpretación* de una oración o conjunto de oraciones y se definió explícitamente la noción más específica de interpretación *oracional*. Se indicó también en ese capítulo que pueden usarse interpretaciones oracionales para probar que no es válido un argumento aleticofuncional o que es consistente un conjunto de premisas. Antes de pasar a las reglas de inferencia aplicables a los cuantificadores existenciales, deseamos definir la noción general de interpretación y usarla para obtener definiciones generales de validez, invalidez e inconsistencia. Se introduce aquí esta discusión, porque resultará útil contar con la noción de interpretación, al discutir las reglas aplicables a los cuantificadores existenciales.

Al dar una interpretación particular de una fórmula u oración, nos recluimos dentro de un *dominio de individuos*, tal como el conjunto de todos los hombres, o el conjunto de los números enteros positivos. Así, dada la oración '$(x)(Hx \rightarrow Mx)$', que aparece en el Ej. 1 de este capítulo, una interpretación del mismo en el dominio de los enteros positivos, que a menudo es llamada una *interpretación*

[7] Los últimos cuatro teoremas corresponden a cuatro de los cinco axiomas usados por Whitehead y Russell (*Principia Mathematica*, primera edición) al presentar su forma axiomatizada de la lógica oracional. (El quinto axioma fue más tarde derivado de éstos por P. Bernays.) Transcritos a la notación lógica del Cap. 1 y usando el hecho de que $-P \vee Q$ es tautológicamente equivalente a $P \rightarrow Q$, estos cuatro teoremas se convierten en

Axioma 1.	$P \vee P \rightarrow P$
Axioma 2.	$P \rightarrow P \vee Q$
Axioma 3.	$P \vee Q \rightarrow Q \vee P$
Axioma 4.	$(P \rightarrow Q) \rightarrow [R \vee P \rightarrow R \vee Q]$

aritmética, es:

$$(x)(x > 3 \rightarrow x > 2).$$

Aquí el predicado '*H*' es reemplazado por el predicado '$>$ 3' (que significa 'es mayor que tres') y el predicado '*M*' por '$>$ 2' (que significa 'es mayor que dos'). Como segundo ejemplo, considérense los dos postulados[8] para preferencia débil dados en el penúltimo ejercicio del apartado anterior:

$$(x)(y)(z)(xQy \;\&\; yQz \rightarrow xQz)$$
$$(x)(y)(xQy \vee yQx).$$

Una simple interpretación aritmética de la conjución de estos dos postulados es la de interpretar el predicado '*Q*' como '$\geq$'. Nótese que si demandamos una interpretación aritmética *verdadera* no podemos interpretar '*Q*' como '$>$', esto es, la preferencia débil no podemos interpretarla como mayor que, pues dados dos números, no se tiene siempre

(1) $$x > y \vee y > x.$$

Obviamente (1) no es válida cuando x y y son el mismo número.

Con los dos ejemplos considerados, no se pretende dar la impresión de que las interpretaciones deben ser siempre de carácter aritmético. Considérese la siguiente oración sobre preferencia estricta, donde, como en el ejercicio a que se acaba de hacer referencia, xPy si y sólo si x es estrictamente preferido a y:

(2) $$(x)(y)(xPy \rightarrow -yPx).$$

Si interpretamos '*P*' como 'es padre de', obtenemos una interpretación verdadera de (2) para el dominio de los seres humanos:

$$(x)(y)[(x \text{ es padre de } y \rightarrow -y \text{ es padre de } x)]$$

Es necesario considerar un ejemplo de una fórmula con variables libres. La técnica de interpretación de esas variables es sencilla: las tres ocurrencias de una variable tal son reemplazadas por el nombre de algún individuo en el dominio de individuos con respecto a la que se hace la interpretación. Así, recurriendo de nuevo a la teoría de la preferencia, considérese

(3) $$xPy \rightarrow (z)(xIz \rightarrow zPy),$$

[8] En toda esta obra usamos las palabras 'axioma' y 'postulado' como sinónimos. Se abandona aquí la vaga distinción entre axiomas y postulados, que se hacen algunas veces en geometría elemental.

que dice:

> Si x es preferida a y, entonces para todos los objetos z, si x y z son indiferentes en cuanto a preferencia (esto es, del mismo valor), entonces z también es preferida a y.

Se encuentra fácilmente una interpretación aritmética verdadera de (3): interprétese la variable libre 'x' como '2', 'y' como '1', 'P' como '>', 'I' como '=', y obtenemos la verdadera aritmética:

(4) $$2 > 1 \rightarrow (z)(2 = z \rightarrow z > 1).$$

Para simplificar la definición de interpretación excluiremos descripciones definidas como 'el esposo de x', o 'el número primo más pequeño mayor que y'. Tenemos entonces como términos: variables, nombres propios y expresiones construidas con variables y nombres usando símbolos de operación, esto es, expresiones como '$2 + 3$', '$x + y$', '$x + 5$'. Para completar nuestra lista de ejemplos intuitivos de interpretaciones, necesitamos indicar cómo se interpretan los nompres propios y los símbolos de operación.

Para nombres propios, la respuesta es sencilla; se manejan justamente como variables libres: todas las incidencias de un nombre propio en una fórmula se reemplazan por el nombre de algún individuo del dominio de individuos con respecto del cual se hace la interpretación. Considérese, por ejemplo, la fórmula

(5) x es una ciudad y x es más grande que San Francisco

Fácilmente se encuentra una interpretación aritmética de (5). Interprétese 'es una ciudad' como 'es un entero positivo', la variable libre 'x' como '1', 'es mayor que' como '>' y el nombre propio 'San Francisco' como '2'. Tenemos entonces como interpretación de (5):

(6) 1 es un entero positivo y $1 > 2$.

La oración (6) es, por supuesto, una interpretación falsa. Si cuantificamos existencialmente la variable 'x' en (5), tenemos

(7) $(\exists\, x)$(x es una ciudad & x es más grande que San Francisco)

Como verdadera interpretación de (7), tenemos entonces:

(8) $(\exists\, x)$(x es un entero positivo & $x > 2$).

Pero es también fácil encontrar una falsa interpretación de (7), a saber:

(9) $(\exists\, x)$(x es un entero positivo & $x < 1$).

En (9) 'San Francisco' es reemplazado por '1' y no por '2' en la interpretación y 'es más grande que' por '<' y no por '>'.

Los símbolos de operación se manejan como predicados; en la interpretación se usa como predicado un símbolo de operación, definido para los individuos del dominio de interpretación. Los símbolos de operación que se dan más frecuentemente en el lenguaje común son probablemente los de adición y multiplicación aritmética. En este capítulo, apenas si trataremos de símbolos de operación, pero a fin de dar una definición completa de interpretación, los mencionaremos de pasada. Como ejemplo de interpretación, podemos interpretar el símbolo de la adición como un símbolo de multiplicación. En este caso el dominio de individuos es igual para la fórmula que para la interpretación dada. Por ejemplo, como una interpretación de

$$(x)(y)(x + y = y + x)$$

tenemos:

$$(x)(y)(xy = yx).$$

Como interpretación de:

$$(\exists y)(x + y > x)$$

tenemos

$$(\exists y)(2y > 2).$$

En este último caso se asigna a la variable 'x' la constante individual '2' en la interpretación.[9]

Sin entrar en detalles minuciosos debe quedar claro que un predicado de una posición se interpreta mediante un predicado de una posición, un símbolo de operación binaria como '+' por un símbolo de operación binaria, y así sucesivamente. Sin embargo, hay cierta flexibilidad en la clasificación de los predicados como de una posición, de dos posiciones, etc., etc. Por ejemplo, podríamos reemplazar el predicado 'es humano' por el predicado aritmético 'es mayor que uno'. que en notación ordinaria parece un predicado de dos posiciones, o hasta podríamos reemplazar 'es humano' por '$(\exists y)$('. . . *es mayor que y*)'. No es necesario para nuestros fines caracterizar exactamente qué predicados pueden ser reemplazados por cuáles otros en una interpretación, pero la idea básica es obvia: el predicado de la interpretación debe tener exactamente el mismo número de posiciones para llenar con variables libres, que el predicado que se interpreta.

[9] En los contextos matemáticos, los nombres propios '1' y '2' se llaman generalmente *constantes individuales* y los predicados *símbolos de relación*.

(No es que en la interpretación, estos lugares se llenen con variables libres. De hecho, es una característica de una interpretación que no tenga variables libres y sea, por lo mismo, verdadera o falsa).

Las observaciones y ejemplos anteriores deben hacer fácil de comprender la definición general de interpretación.

> *La oración* **P** *es una interpretación de la fórmula* **Q** *con respecto del dominio de individuos* D *si y sólo si* **P** *puede obtenerse de* **Q**, *colocando predicados y símbolos de operación definidos para el dominio* D, *en vez de los predicados y símbolos de operación, respectivamente, de* **Q**, *y colocando nombres propios de individuos de* D *en vez de nombres propios* (*esto es, constantes individuales*) *y variables libres de* **Q**.[10]

Usando esta noción de interpretación podemos ahora definir la *validez universal*, la *consecuencia lógica* o *implicación lógica* y la *consistencia*.

> *Una fórmula es universalmente válida si y sólo si es verdadera toda interpretación de la misma en todo dominio de individuos no vacío.*

La idea intuitiva en que se basa esta definición es que las fórmulas universalmente válidas deben ser verdaderas en todo mundo posible. Si su verdad se basa en hechos contingentes del mundo tempoespacial, no tienen una genuina validez universal.

> *Una fórmula* **Q** *se sigue lógicamente de una fórmula* **P** *si y sólo si en todo dominio de individuos no vacío, toda interpretación que haga verdadera a* **P** *también hace verdadera a* **Q**.

Una definición equivalente es simplemente:

> **Q** *se sigue lógicamente de* **P** *si y sólo si la condicional* $\mathbf{P} \rightarrow \mathbf{Q}$ *es universalmente válida.*

Una segunda idea intuitiva en que se apoyan estas definiciones es que una fórmula se sigue lógicamente de otra, justamente, en virtud de sus respectivas formas lógicas. Tomar en cuenta todas las interpretaciones de una fórmula equivale a prescindir de todo, excepto su pura estructura lógica. Debe hacerse notar que al llevar a término este proceso de abstracción se niega la validez universal de ciertas oraciones del lenguaje ordinario que parecerían ser verdaderas en todo mundo posible. Un ejemplo típico es 'todos los solteros no son

[10] Una definición completamente precisa requiere los conceptos sobre teoría de conjuntos desarrollados en la Parte II de este libro. Esta definición se extiende o amplía ligeramente en el § 4.3 para justificar el dispositivo técnico usado en el manejo de cuantificadores existenciales.

casados. El análisis lógico de esas oraciones es un asunto complicado y sutil que evitaremos en este libro. Afortunadamente, el desarrollo deductivo sistemático de cualquier rama de las matemáticas o ciencia teórica puede proceder sin recurso explícito a oraciones cuya verdad se sigue únicamente de los significados de los predicados utilizados.[11]

La definición de consistencia fácilmente se deduce del Cap. 2.

> *Una fórmula es consistente si y sólo si tiene cuando menos una interpretación verdadera en algún dominio de individuos no vacío.*

(Al manejar un conjunto de premisas, podemos proceder a su conjunción para obtener una sola fórmula).

En las tres definiciones que se acaban de dar sólo se han tomado en cuenta dominios de individuos *no vacíos*. Posiblemente el argumento único más convincente en pro de esta restricción es que necesitamos considerar como lógicamente válida la inferencia de

$$(x)Hx$$

a:

$$(\exists x)Hx$$

Y, correspondientemente, necesitamos que resulte inconsistente la fórmula

$$(x)Hx \ \& -(\exists x)Hx$$

Si se incluye el dominio vacío, ninguna de estas necesidades quedará satisfecha.

Para mayor claridad, es también deseable definir la noción de *argumento inválido*.

> *Un argumento es inválido si y sólo si hay una interpretación en algún dominio no vacío que haga verdaderas a sus premisas y falsa a su conclusión.*

Podemos aplicar esta definición para demostrar explícitamente que es necesaria la restricción de imprecisión impuesta a la generalización universal, que se introdujo en el apartado anterior. Nuestra notación para las deducciones es de tal naturaleza, que una línea de una deducción debe seguirse lógicamente del conjunto de premisas correspondiente al conjunto de números situados a la izquierda de la línea. Ahora bien, dada la premisa 'Hx', supóngase que pudimos deducir '$(x)Hx$'. Es fácil encontrar una interpretación aritmética para la cual 'Hx' sea verdadera y '$(x)Hx$' sea falsa. Interprétese la variable libre

[11] Casi todo el Cap. 12 tiene que ver con este particular; sólo es necesaria una somera teoría del significado para aplicar el método axiomático a una rama dada de la ciencia.

'*x*' como '1' y '*H*' como el predicado aritmético 'es menor que 2'. Obviamente 1 es menor que 2 y de aquí que la premisa sea verdadera en esta interpretación, pero ciertamente es falso que todo entero positivo sea menor que 2, de donde la conclusión es falsa en esta interpretación y el argumento no es válido (es inválido). (No debe resultar ninguna confusión del hecho de que en este ejemplo sea libre la '*x*' de la premisa y en cambio sean ligadas ambas incidencias de '*x*' en la conclusión).

Encontrar una interpretación que muestre que un argumento es lógicamente inválido equivale a encontrar un *contraejemplo* del argumento. La habilidad para construir contraejemplos adecuados de argumentos inválidos es casi tan importante como la habilidad de construir deducciones lógicas a partir de argumentos válidos.

En el Cap. 2 vimos que hay un método mecánico (mediante el uso de tablas de verdad) para probar la validez o invalidez aleticofuncional de un argumento. Ese método mecánico es llamado a menudo un *procedimiento de decisión*. En un sentido, la existencia de un procedimiento de decisión para los argumentos aleticofuncionales trivializa* el tema. Afortunada o desafortunadamente, no es posible esa trivialización de la lógica de la cuantificación. Fue rigurosamente probado en 1936 por el lógico americano contemporáneo Alonso Church que en la lógica de predicados de primer orden[12] no hay procedimientos de decisión, es decir, no hay prueba mecánica de la validez de las fórmulas arbitrarias. Puesto que todas las matemáticas pueden ser formalizadas dentro de la lógica de predicados de primer orden,[13] la existencia de ese procedimiento de decisión tendría consecuencias asombrosas; podría construirse una máquina para resolver cualquier problema matemático o para decidir sobre la validez o invalidez de cualquier argumento matemático. Pero el teorema de Church arruina de una plumada todos los sueños de los estudiantes de lógica y de matemáticas. No solamente no hay un procedimiento de decisión; su teorema establece que nunca lo habrá.[14]

---

* El neologismo es necesario y, además, adecuado, por cuanto está de acuerdo con los procedimientos lexicológicos del castellano. Téngase en cuenta que la lógica simbólica es materia en gran parte no explorada en castellano (en cuanto a su formulación verbal), que representa la confluencia del pensamiento filosófico, matemático y lingüístico y que, por lo mismo, demanda la creación del neologismo. (N. del T.)

[12] La lógica de predicados de primer orden es la lógica de los conectivos oracionales y cuantificadores de variables individuales, esto es, la lógica de las fórmulas definidas en el Cap. 3. "Primer orden" quiere decir que no se permite la cuantificación de los predicados.

[13] Los desarrollos tipo de la teoría axiomática de los conjuntos tienen como una de sus metas establecer este hecho en detalle.

[14] Existen procedimientos de decisión para varios dominios especiales de las matemáticas Probablemente el resultado positivo más importante es el procedimiento de decisión de Alfredo Tarski para el álgebra elemental de los números reales, publicado por

Por otra parte, puede demostrarse que son completas las reglas de inferencia dadas en este capítulo, juntamente con las del Cap. 2, lo que significa que si **Q** se sigue lógicamente de **P**, entonces **Q** es deducible de **P** por medio de nuestras reglas de inferencia. La primera prueba de integridad para un conjunto de reglas de inferencia de la lógica de predicados de primer orden fue dada por Kurt Gödel en 1930.

Aunque nuestra lógica básica es completa, no lo son la mayor parte de las teorías de alguna complejidad, en el sentido de que no es posible dar una lista de axiomas o postulados de la teoría de los cuales sea posible deducir todas las otras aseveraciones verdaderas de la teoría. En particular, fue demostrado en 1931 por Gödel que la teoría elemental de los enteros positivos, y a fortiori cualquier teoría que incluya teoría elemental de los números, es incompleta en el sentido que se acaba de indicar. El teorema de Gödel sobre la deficiencia o no integridad de la teoría elemental de los números es probablemente el teorema más importante en los anales de la lógica moderna.

Los últimos tres párrafos han tratado de indicar a grandes rasgos la clase de importantes nociones y resultados que se apiñan alrededor del concepto de validez lógica. Pero estos párrafos son superficiales y constituyen una digresión de nuestra empresa principal.[15]

Antes de volver a nuestras reglas de inferencia debe mencionarse otra aplicación útil del método de interpretación, a saber, su uso para demostrar que una fórmula es lógicamente independiente de un conjunto de otras fórmulas. La aplicación es fácil: mediante el uso de una interpretación se demuestra que la fórmula dada no es una consecuencia lógica del conjunto dado de fórmulas. La aplicación más común de este método es al problema de demostrar que los axiomas de un conjunto en que se basa una teoría, son mutuamente independientes, esto es, dos a dos. La técnica consiste en dar para cada axioma una interpretación tal que el axioma dado sea falso y los restantes sean verdaderos en la interpretación. Si el axioma dado pudiera ser deducido de los restantes, resultaría una violación de nuestra definición básica de consecuencia lógica.

---

primera vez en 1948. Dada una oración sobre números reales, construida con variables individuales o con constantes individuales que denoten números reales particulares, con la relación '$<$', los símbolos operacionales de adición y multiplicación, el signo de igualdad, los conectivos oraciones y con cuantificadores, Tarski da una prueba mecánica para decidir sobre la verdad o falsedad de la oración.

[15] En el muy completo tratado de S. C. Kleene, *Introduction to Metamathematics* Nueva York, 1952, pueden encontrarse detalles relativos a estos teoremas.

Para ilustrar el método de probar la independencia de los axiomas, podemos considerar de nuevo los dos axiomas de la preferencia débil:

I. $(x)(y)(z)(xQy \mathbin{\&} yQz \rightarrow xQz)$
II. $(x)(y)(xQy \vee yQx)$.

Se prueba la independencia del Axioma 1 por medio de la siguiente interpretación aritmética:

$$xQy \text{ si y sólo si } x \leq y + 1.$$

Para cualesquiera dos enteros positivos x y y es verdad que

$$x \leq y + 1 \quad \text{o} \quad y \leq x + 1,$$

y en esta interpretación resulta verdadero el Axioma II; pero el Axioma I no es verdadero para todos los enteros x, y y z, pues aunque

$$3 \leq 2 + 1,$$

y

$$2 \leq 1 + 1,$$

es falso que

$$3 \leq 1 + 1.$$

Esto es, si en esta interpretación del Axioma I aplicamos la especificación universal, haciendo 'x' igual a '3', 'y' igual a '2' y 'z' igual a '1', obtenemos una oración falsa. Ya ha sido mencionada la interpretación aritmética que establece la independencia del Axioma II: interprétese el predicado 'Q' como 'es mayor que'.

En virtud de que el método de interpretación es el único método general para demostrar la invalidez de los argumentos, la consistencia de las premisas o la independencia de los postulados, será útil una lista de reglas explícitas para formular interpretaciones. En estas reglas no se incluirá ninguna idea que no se haya expuesto ya; su propósito es el de llamar la atención hacia los errores más comunes en la aplicación del método de interpretación.

La primera regla simplemente sugiere un formato tipo (estándar) para usarlo en las aplicaciones:

REGLA I. *Al aplicar el método de interpretación, siempre y con toda claridad:*

(i) *Enúnciese qué conjunto de objetos es el dominio de interpretación;*

(ii) *Enúnciese, en forma de equivalencias o identidades, la interpretación de predicados, nombres propios, símbolos de operación y variables libres;*

(iii) *Anótese la interpretación de cada oración original y enúnciese si es verdadera o falsa.*

Sea, por ejemplo, el argumento:

> Algunos hombres son mentirosos.
> Adams es un hombre.
> En consecuencia, Adams es un mentiroso.

Podemos representarlo simbólicamente en esta forma:

$$\frac{\begin{array}{l}(\exists x)(Mx \,\&\, Lx)\\ Ma\end{array}}{La}$$

Demostramos que este argumento es inválido, haciendo la siguiente interpretación.

(i) Dominio de interpretación = conjunto de los enteros positivos.

(ii) Interpretación de predicados y nombres propios:

$Mx \leftrightarrow x$ es un entero positivo
$Lx \leftrightarrow x$ es un entero par
$a = 1.$

(iii) Interpretación de oraciones originales:

T (1) ($\exists$ x)(x es un entero positivo y x es un entero par)
T (2) 1 es un entero positivo
F (3) En consecuencia, 1 es un entero par.

De las más conocidas verdades de la aritmética resulta inmediatamente evidente que (1) y (2) son verdaderas, mientras que (3) es falsa, lo cual establece que el argumento no es válido.

> REGLA II. *Las variables ligadas de la oración original deben quedar intactas en la interpretación de la oración. En particular, no se cambia ningún cuantificador.*

Para ilustrar la violación de esta regla, sea el argumento:

> Hay mentirosos.
> En consecuencia, hay ladrones.

Representamos esto simbólicamente como sigue:

$$\frac{(\exists x)(Lx)}{(\exists x)(Tx)}.$$

Usamos la siguiente interpretación para establecer la obvia invalidez del argumento.

(i) Dominio de interpretación = {1, 3}.
(ii) Interpretación de predicados:

$$Lx \leftrightarrow x = 1$$
$$Tx \leftrightarrow x = 2$$

(iii) Interpretación de oraciones originales:

$$\frac{\text{T} \quad (1)\ (\exists x)(x = 1)}{\text{F} \quad (2)\ (\exists x)(x = 2)}.$$

Claramente, para el dominio {1, 3}, la oración (1) es verdadera y la (2) falsa. Quizá a causa de que el dominio es un conjunto finito, muchos estudiantes al principio tratan de ofrecer algo como

$$(1')\ 1 = 1$$

como interpretación de (1), por ejemplo. Pero (1') viola rotundamente la Regla II, pues no contiene un cuantificador existencial que use la variable '*x*', como sí lo contiene la oración original. También sería un error ofrecer como interpretación

$$\frac{(1'')\ x = 1}{(2'')\ x = 2}.$$

Aquí, los cuantificadores se omiten incorrectamente. Además, ya se ha hecho notar que una interpretación no debe contener variables libres, pues debe ser una oración verdadera o falsa. Este último punto puede ser enunciado como la regla siguiente.

Regla III. *La interpretación de una oración o fórmula no debe contener variables libres.*

De las reglas segunda y tercera se sigue la consecuencia de que es indebido especificar universal o existencialmente al dar una interpretación. La especificación nunca debe confundirse con la interpretación. Sin embargo, al establecer la falsedad de una oración interpretada que tiene un cuantificador universal, puede ser útil la aplicación de la especificación universal y al establecer la verdad de una oración interpretada que tiene un cuantificador existencial, el método más directo puede consistir en exhibir un objeto (del dominio de interpretación), que satisfaga la fórmula existencialmente cuantificada. Por ejemplo, para demostrar que la anterior oración (1), [esto es, '$(\exists x)\ (x = 1)$'], es verdadera, es suficiente encontrar un objeto del dominio que satisfaga la fórmula '$x = 1$'. Podemos indicar esta selección por una fórmula como (1'). De modo semejante,

supóngase que tenemos como interpretación de la conclusión de un argumento inválido, la oración.

(3) $$(x)(x > 1 \rightarrow x > 3)$$

donde el dominio de interpretación es el conjunto de los enteros positivos. A fin de que (3) sea verdadera, debe serlo para cualquier caso específico de un entero positivo, de manera que para demostrar que es falsa necesitamos exhibir sólo un entero positivo para el cual no sea verdadera. En este caso tomamos $x = 2$, esto es, especificamos a 'x' sustituyéndolo por dos, y obtenemos la falsa oración:

(4) $$2 > 1 \rightarrow 2 > 3.$$

Es necesario insistir de nuevo en que tanto (1') como (4) no son interpretaciones, sino meras oraciones auxiliares para establecer la verdad o falsedad, según sea el caso, de las interpretaciones (1) y (3), respectivamente.

Las reglas cuarta y quinta se refieren a la clase de predicados que pueden usarse como interpretaciones de un predicado dado.

REGLA IV. *La interpretación de un predicado debe usar el mismo número de variables distintas que el predicado original.*

Supóngase, por ejemplo, que deseábamos demostrar que son consistentes las siguientes premisas:

$$(\exists x)(y)(xQy)$$
$$(x)(y)(xQy \rightarrow yQx)$$
$$(x)(y)(z)(xQy \ \& \ yQz \rightarrow xQz).$$

A fin de demostrar que estas tres premisas son consistentes, necesitamos encontrar una interpretación para la cual sean todas verdaderas. (Aunque estas tres premisas usan el predicado 'Q' de preferencia débil, analizamos su consistencia independientemente de los dos axiomas sobre la preferencia dados anteriormente). Aquí resulta procedente una interpretación con un solo elemento de su dominio.

(i) Dominio de interpretación = {1}.

(ii) Interpretación del predicado:

$$xQy \leftrightarrow x = y.$$

(iii) Interpretación de las oraciones originales:

T (1) $(\exists x)(y)(x = y)$

T (2) $(x)(y)(x = y \rightarrow y = x)$

T (3) $(x)(y)(z)(x = y \ \& \ y = z \rightarrow x = z)$.

Para el dominio {1}, los únicos casos de especificación universal de (2) y (3), son, respectivamente:

(2′) $$1 = 1 \rightarrow 1 = 1$$

(3′) $$1 = 1 \,\&\, 1 = 1 \rightarrow 1 = 1$$

que son obviamente verdaderas. En el caso de (1) debemos tomar $x = 1$, y obtener:

(4′) $$(y)(1 = y),$$

que es verdadera para la única especificación posible de 'y'. (Notamos de nuevo que (2′), (3′) y (4′) *no* son interpretaciones de las premisas dadas.)

Hubiera sido una violación de la Regla IV tomar como interpretación de 'Q':

(5) $$xQy \leftrightarrow x = x,$$

pues la fórmula $x = x'$ solamente usa una variable, mientras que '$x'Qy$' usa dos. Naturalmente la interpretación de las premisas derivadas de (5) sería tal, que las tres interpretaciones resultantes serían todas verdaderas, pero la forma de las interpretaciones sería incorrecta.[16]

Regla V. *La interpretación de un predicado es igual en todas sus incidencias, sin importar qué variables se usen con el predicado.*

Así, dada la premisa:

$$(\exists x)(Hx \,\&\, (y)(Hy \rightarrow Lxy))$$

Sería un error interpretarla como:

$$Hx \leftrightarrow x > 1$$
$$Hy \leftrightarrow y > 3.$$

Violando la Regla V es fácil dar una demostración espúrea de que

[16] De hecho la Regla IV puede ser debilitada hasta darle la formulación. Un predicado *no debe usar más* variables distintas que el predicado original (véase la discusión relacionada de las definiciones, Pág. 203). Por otra parte, cuando la interpretación de un predicado usa variables ligadas, como por ejemplo $xQy \leftrightarrow (\exists z)(x \geq z \,\&\, z \geq y)$, entonces debe modificarse la Regla IV, para que establezca que una interpretación de un predicado no debe usar más variables distintas libres que el predicado original. Además, debe tomarse en cuenta una restricción a las variables ligadas de la interpretación, a saber, que un cuantificador de la interpretación no debe capturar variables de la fórmula original. Sin embargo, ninguna de las aplicaciones consideradas en la secuela requiere variables ligadas en la intepretación de un predicado.

un argumento válido es inválido. Por ejemplo, considérese el argumento válido:

$$\frac{(\exists x)Fx}{(\exists y)Fy}.$$

Sea el conjunto de los números enteros positivos el dominio e interprétese, incorrectamente, como:[17]

$$Fx \leftrightarrow x = 1$$
$$Fy \leftrightarrow y \neq y.$$

Puesto que la oración '$(\exists x)(x = 1)$' es verdadera y la oración '$(\exists y)(y \neq y)$', es falsa, concluiríamos erróneamente, basados en la interpretación equivocada, que el argumento es inválido.

Además, al hacer la interpretación de un predicado, no es necesario asentar equivalencias de todas las distintas variables que se dan con este predicado. Un ejemplo es suficiente para indicar la interpretación del predicado; a continuación se hace una adecuada sustitución de las variables al interpretar las fórmulas dadas. Por ejemplo, al demostrar que las tres premisas que siguen de la Regla IV son consistentes, dimos como única interpretación de 'Q':

$$xQy \leftrightarrow x = y,$$

y a continuación, al interpretar las premisas (2) y (3) como (2') y (3') respectivamente, hicimos los cambios obvios y necesarios de variables; esto es, reemplazamos

'yQx' por '$y = x$'
'yQz' por '$y = z$'
'xQz' por '$x = z$'.

Regla VI. *La interpretación de nombres propios y variables libres debe designar objetos del dominio de interpretación.*

Considérense, por ejemplo, tres posibles postulados de preferencia donde 'a' es el nombre usado para algún objeto, digamos, una primera edición original de la *Crítica de la Razón Pura* de Kant:

(1) $(\exists x)(xPa)$

(2) $(x)(y)(xPy \rightarrow -(yPx))$

(3) $(x)(y)(z)(xPy \;\&\; yPz \rightarrow xPz)$.

Deseamos demostrar por el método de interpretación que el postulado (1) es independiente de los otros dos. Como se ha explicado anterior-

[17] El predicado '$\neq$' se lee 'no idéntico con'. Podríamos simbolizar '$y \neq y$' con '$-(y = y)$'.

mente, establecer esa independencia equivale a demostrar que (1) no es una consecuencia lógica de (2) y (3). Consecuentemente, deseamos encontrar una interpretación para la cual (1) sea falsa y (2) y (3) verdaderas. Supóngase que tomamos como dominio de interpretación el conjunto de enteros positivos. Sería entonces una violación de la Regla VI interpretar '*a*' como un objeto, digamos, −2, que no está en el dominio. Una interpretación correcta es la siguiente:

(i) Dominio de interpretación = conjunto de enteros positivos
(ii) Interpretación de predicado y nombre propio:

$$xPy \leftrightarrow x < y$$

$$a = 1$$

(iii) Interpretación de postulados:

F (1′) $(\exists x)(x < 1)$
T (2′) $(x)(y)(x < y \rightarrow -(y < x))$
T (3′) $(x)(y)(z)(x < y \,\&\, y < z \rightarrow x < z)$.

Puesto que '−2' como interpretación de '*a*' da también una falsa interpretación de (1) para este dominio:

$$(1'')\ (\exists x)(x < -2),$$

podría pensarse que la restricción expresada por la Regla VI no es realmente necesaria. Sin embargo, su violación puede conducir a la demostración falaz de que es inválido un argumento válido o de que es independiente un postulado dependiente. Por ejemplo, considérense los dos postulados:

$$Fa$$

$$(x)Fx.$$

Está claro que el primero puede ser inmediatamente deducido del segundo por especificación universal. Sin embargo, si tomamos como nuestro dominio de interpretación el conjunto de enteros positivos, interpretamos '*F*' como:

$$Fx \leftrightarrow x \geq 1$$

y '*a*' como '−3', obtenemos entonces el resultado falaz de que el primer postulado es independiente del segundo, puesto que '$-3 \geq 1$' es falso, mientras que '$(x)(x \geq 1)$' es verdadero para el dominio escogido.

Estrechamente relacionada con la Regla VI, existe una regla co-

rrespondiente para los signos de operación, que será útil principalmente en los capítulos subsecuentes.

REGLA VII. *La interpretación de los símbolos de operación debe ser tal que la interpretación de cualquier término que use el símbolo de operación se refiera a un elemento del dominio de interpretación.*

Así, dada la oración:

$$(x)(y)(x + y = y + x)$$

será un error fijar el conjunto {1} como dominio de interpretación e interpretar el símbolo de operación '+' como adición aritmética, pues entonces el término '1 + 1' designa un objeto que no pertenece al dominio de interpretación. Una forma técnica de enunciar el requisito establecido por la Regla VII, es decir que la interpretación de un símbolo de operación debe tener la *propiedad de oclusión* con respecto al dominio de interpretación.

Como segundo ejemplo, supóngase que se nos dan los siguientes dos postulados de un símbolo 'o' de operación binaria:

(1) $$(x)(y)(x \circ y = y \circ x)$$

(2) $$(\exists x)(y)(y \circ x = y),$$

y deseamos probar que el primer postulado es independiente del segundo. Un error común en relación con un ejemplo de este tipo es el de tomar como dominio de interpretación el conjunto {1, 2} e interpretar el símbolo de operación como signo de división:

(3) $$x \circ y = x/y.$$

Entonces (2) parece verdadera, pues podemos hacer $x = 1$ y para cualquier número y

$$y/1 = y.$$

Y (1) parece falsa, pues especificando universalmente y haciendo 'x' igual a '1' y 'y' igual a '2', obtenemos la falsedad obvia:

$$\tfrac{1}{2} = \tfrac{2}{1}.$$

Sin embargo, la especificación hecha en (1) muestra que, tomados juntamente el dominio {1, 2} y la interpretación (3), se viola la Regla VII, pues el número 1/2 no pertenece al conjunto {1, 2}. Para indicar por qué la restricción impuesta por la Regla VII no es caprichosa, sino esencial, necesitamos anticipar algunas de las ideas expuestas al principio del capítulo siguiente. Tomamos allí como una

verdad lógica que no depende de ninguna premisa, la identidad **t** = **t**, donde **t** es cualquier término. Por ejemplo, en el presente contexto, si admitimos a '1/2' como término, entonces la oración:

(4) $$\frac{1}{2} = \frac{1}{2}$$

es verdadera, pero de (4) podemos inferir inmediatamente la oración

$$(\exists x)(\tfrac{1}{2} = x)$$

que es obviamente falsa cuando el dominio de interpretación es {1, 2}.

En conexión con cierto número de ejercicios de apartados subsecuentes, puede ser útil exhibir dos interpretaciones diferentes, una con un dominio finito y otra con un dominio infinito de interpretación, que prueban que (1) es independiente de (2). Si interpretamos como de división el símbolo de operación '$\circ$', necesitamos un dominio infinito: el conjunto de números positivos racionales. (Un número positivo racional es un número positivo que es la razón de dos enteros positivos.) La Regla VII queda satisfecha con este dominio de interpretación de '$\circ$'; además, la interpretación resultante de (1) es falsa y la de (2) es verdádera. Como segunda interpretación, sea el dominio de interpretación el conjunto {1, 2}, e interprétese '$\circ$' por medio de la identidad:

(5) $$x \circ y = x.$$

Cuando el dominio es finito se da algunas veces una interpretación como (5) por medio de una tabla o *matriz* más bien que con una identidad. De este modo podríamos reemplazar (5) por la matriz:

$\circ$	1	2
1	1	1
2	2	2

Esta tabla o matriz se usa en la siguiente forma. Para encontrar qué elemento es, digamos 1 o 2, vemos la anotación correspondiente al primer renglón y a la segunda columna y encontramos:

$$1 \circ 2 = 1.$$

Notamos, por otra parte, que

$$2 \circ 1 = 2.$$

Así, conforme a esta interpretación es falso el Postulado (1), puesto que la especificación de '1' para 'x' y '2' para 'y' da el absurdo de que $1 = 2$. En cambio, el Postulado (2) es, como se desea, verdadero en esta interpretación, como puede verse haciendo $x = 1$, puesto que la oración:

$$(y)(y \circ 1 = y)$$

es aquí verdadera.

Para terminar este apartado, resumimos de nuevo las tres más importantes aplicaciones del método de interpretación:

Prueba de invalidez de los argumentos.
Prueba de consistencia de las premisas.
Prueba de independencia de los axiomas.

EJERCICIOS

1. Por medio de interpretaciones aritméticas pruebe que las siguientes fórmulas no son universalmente válidas.

(a) $(x)(Hx \,\&\, Ax \rightarrow Mx)$
(b) $(\exists x)Hx$
(c) $(\exists x)(Hx \vee -Ax)$
(d) $(x)(Hx \,\&\, -Mx)$

2. Dé interpretaciones (aritméticas o de otra naturaleza) para probar que no son válidos los siguientes argumentos. (Nótese que todas las conclusiones son verdaderas, pero no son válidas.)

(a) Todos los hombres son animales. Todos los hombres son mortales. En consecuencia, todos los animales son mortales. (Ax, Hx, Mx)
(b) Nueva York está al norte de Washington. Boston está al norte de Nueva York. En consecuencia, Boston está al norte de Washington. (Nxy, n, w, b)
(c) Algunos marineros son ignorantes. Algunos americanos son ignorantes. En consecuencia, algunos marinos son americanos. (Ax, Ix, Ex)
(d) Todos los hombres son animales. Algunos animales son de corta vida. En consecuencia, algunos hombres son de corta vida. (Sx)

3. Formule interpretaciones para probar que los siguientes conjuntos de premisas son consistentes. Los primeros tres ejemplos tratan de la teoría de la preferencia, pero son independientes de suposiciones anteriores sobre preferencias e indiferencia.[1*]

(a) $(\exists x)(y)(xQy)$
$(x)(y)(\exists z)(xQz \,\&\, zQy)$
(b) $(x)(y)(xPy \rightarrow -yPx)$
$(x)(y)(xIy)$
(c) $(x)(\exists y)(yPx)$
$(x)(y)(xPy \rightarrow -yPx)$
$(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$
(d) Todos los unicornios son animales. Ningún unicornio es animal. (Ux)

[1*] Las propiedades de orden que se tratan en estos ejercicios sobre la teoría de la preferencia son de importancia en muchos dominios de la ciencia. En el Cap. 10 hay una extensa discusión de las relaciones de ordenación.

4. Pruebe por el método de interpretación que los postulados son mutuamente independientes en como uno de los siguientes casos. Sugestión: al usar interpretaciones aritméticas es, a menudo, conveniente restringir el dominio de individuos a algún conjunto finito de enteros como los dos primeros o los diez primeros. De nuevo se toman ejemplos de la teoría de la preferencia.

(a) $(x)(y)(xPy \rightarrow -yPx)$
$(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$
(b) $(x)(y)(z)(xIy \,\&\, yIz \rightarrow xIz)$
$(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$
(c) $(x)(\exists y)(xPy)$
$(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$
(d) $(x)(y)(xIy \rightarrow yIx)$
$(x)(y)(xPy \rightarrow -(xIy \vee yPx))$

5. Formule interpretaciones para probar que son consistentes los siguientes conjuntos de postulados, que tratan de un símbolo binario de operación. Tenga cuidado de no violar la Regla VII.

(a) $(x)(y)(x \circ y = y \circ x)$
$(\exists z)(x)(y)(x \circ y = z)$
(b) $(x)(\exists y)(x \circ y = 1)$.
$(x)(\exists z)(x \circ z = 0)$
$(x)(\exists w)(x \circ 1 = w)$
(c) $(\exists x)(\exists y)(x \circ y \neq y \circ x)$
$(x)(\exists y)(x \circ y = 0)$
$(x)(\exists z)(z \circ x = 0)$

§4.3. Inferencias restringidas con cuantificadores existenciales.[19]

Al estudiar el modo de formular las reglas de inferencia que rigen a los cuantificadores existenciales, probablemente la idea más natural es la de proceder como hicimos para los cuantificadores universales, e introducir una regla de *especificación existencial* que nos permita simplemente omitir un cuantificador existencial. Por ejemplo, de:

(1) $(\exists x)Hx$

inferiríamos:

(2) $Hx.$

Sin embargo, un momento de reflexión muestra que esta regla, si se usa irrestrictamente, produce inferencias inválidas. Por ejemplo, si interpretamos '*H*' como 'mayor que 1' y sustituimos en (2) la variable '*x*' por el nombre '1', en esta interpretación aritmética, de la verdad de que hay un número mayor que 1, inferimos la oración falsa 1 es mayor que 1. Por otra parte, si generalizamos universalmente la oración (2), podríamos inferir '$(x)Hx$' de '$(\exists x)Hx$', que es claramente inválida.

[19] La restricción impuesta en este apartado es que las fórmulas cuantificadas existencialmente deben contener sólo una variable individual.

Se han desarrollado algunos métodos para evitar esas inferencias inválidas.[20] El dispositivo técnico aquí utilizado, aunque aparentemente nuevo en los anales de la lógica de predicados de primer orden, se aproxima al método usado continuamente en matemáticas en las pruebas intuitivas. El dispositivo consiste en reemplazar un cuantificador existencial por una constante 'temporal'; esto es, se introduce un símbolo que funciona como nombre de los propósitos que se tienen a la vista. Este artificio se justifica por el hecho de que el cuantificador existencial garantiza que se puede considerar que un individuo está representado por la nueva constante individual. Por ejemplo, supóngase que deseamos deducir 'Hay un mortal', esto es, '$(\exists x)Mx$' de las premisas '$(x)(Hx \to Mx)$' y '$(\exists x)Hx$'. Podríamos proceder como sigue:

{1}	(1) $(x)(Hx \to Mx)$	P
{2}	(2) $(\exists x)Hx$	P
{2}	(3) $H_{\text{John Doe}}$	Nombre ambiguo aplicado conforme a (2), por especificación existencial
{1}	(4) $H_{\text{John Doe}} \to M_{\text{John Doe}}$	1 *US*
{1, 2}	(5) $M_{\text{John Doe}}$	3, 4 T
{1, 2}	(6) $(\exists x)Mx$	Elimínese de (5) el nombre ambiguo por generalización existencial

El uso del nombre ambiguo 'John Doe' en la deducción anterior nos permite eliminar el cuantificador existencial y aplicar métodos oracionales en la forma normal. El nombre ambiguo 'John Doe' es adecuado en argumentos relativos a seres humanos. En contextos matemáticos se introduce a menudo un nombre ambiguo fijando un asterisco a una variable, por ejemplo, 'x^*' o 'z^*'. Naturalmente, no nos referimos a ningún individuo en particular cuando usamos 'John Doe', y puede con razón alegarse que 'John Doe' no es, propiamente, un nombre propio; por eso usamos la expresión 'nombre ambiguo'. La premisa existencial '$(\exists x)Hx$' garantiza que hay algún individuo a quien podamos atribuir el nombre 'John Doe'.

A fin de establecer un procedimiento uniforme con respecto de

[20] Véase, por ejemplo, W. V. Quine, *Methods of Logic*, Nueva York, 1950, e I. Copi, *Symbolic Logic*, Nueva York, 1954. En este particular debe advertirse que la noción de *imprecisión* del sistema de Quine es enteramente diferente de la desarrollada en el primer apartado de este capítulo.

los nombres ambiguos, usamos para este fin letras griegas minúsculas:

$$\alpha, \beta, \gamma, \delta, \alpha_1, \beta_1, \gamma_1, \delta_1 \ldots$$

Como lo indica la rápida deducción anterior tenemos dos reglas que rigen a los cuantificadores esenciales y a los nombres ambiguos:

ESPECIFICACIÓN EXISTENCIAL (*ES*). *La aserción de que hay algo que satisface una condición dada implica la aserción de que esta condición dada es satisfecha por algún individuo a quien se puede aplicar un nombre.*

GENERALIZACIÓN EXISTENCIAL (*EG*). *La aserción de que una condición es satisfecha por un individuo designado por un nombre, implica la aserción de que esta condición es satisfecha por algún individuo.*

Usando una letra griega como símbolo de un nombre ambiguo e introduciendo las abreviaturas *ES* y *EG*, podemos formular de nuevo como sigue la anterior deducción:

EJEMPLO 4.

{1}	(1) $(x)(Hx \rightarrow Mx)$	P
{2}	(2) $(\exists x)Hx$	P
{2}	(3) $H\alpha$	2 *ES*
{1}	(4) $H\alpha \rightarrow M\alpha$	1 *US*
{1, 2}	(5) $M\alpha$	3, 4 T
{1, 2}	(6) $(\exists x)Mx$	5 *EG*

Antes de enunciar formalmente las reglas de especificación y generalización existenciales, conviene indicar por qué es preciso imponer cierta restricción. Desde las premisas verdaderas '$(\exists x)Hx$' y '$(\exists x)-Hx$', podemos deducir la conclusión falsa '$(\exists x)(Hx \,\&\, -Hx)$', si no se restringe el uso de nombres ambiguos.

{1}	(1) $(\exists x)Hx$	P
{2}	(2) $(\exists x)-Hx$	P
{1}	(3) $H\alpha$	1 *ES*
{2}	(4) $-H\alpha$	2 *ES* (falazmente)
{1, 2}	(5) $H\alpha \,\&\, -H\alpha$	3, 4 T
{1, 2}	(6) $(\exists x)(Hx \,\&\, -Hx)$	5 *EG*

La dificultad se suscita del hecho de que los nombres ambiguos, como todos los nombres, no pueden ser usados indiscriminadamente. Se da perfecta cuenta de esto aquel que da a una persona amada el nombre de otra persona, *anteriormente* amada. Habiéndose escogido

en la línea (3) a 'α' como nombre de alguna persona humana postulada en la línea (1), no podemos en la línea (4) usar también 'α' como nombre de algo que no es humano, como lo postula la línea (2). Ese alegre proceso de dar nombres tiene que conducir a un error, tal como podríamos inferir una conclusión falsa de hechos verdaderos relativos a dos individuos llamados 'Pedro Pérez', si no diseñamos de algún modo un dispositivo de notación, para distinguir a qué Pedro Pérez se refería una expresión dada. La restricción que ahora imponemos para poner coto a esos argumentos inválidos consiste en exigir que cuando se introduzca en una deducción por especificación existencial un nombre ambiguo, ese nombre no haya sido previamente usado en la deducción. Esta es la restricción más simple, aunque no la más débil que podemos enunciar. Requiere una aplicación inmediata de la especificación existencial, antes de la especificación universal. Así, en el Ej. 4 anterior es necesario que la línea (3) preceda a la línea (4). Si la línea (4) hubiera ido primero, 'α' habría sido ya introducido y entonces no habría podido introducirse por especificación existencial.

El enunciado formal de ambas reglas es como sigue:

REGLA DE ESPECIFICACIÓN EXISTENCIAL: *ES. Si una fórmula* **S** *resulta de una fórmula* **R** *sustituyendo todas las incidencias libres de una variable* **v** *en* **R** *por un nombre ambiguo que no haya sido previamente usado en la deducción, entonces* **S** *es deducible de* $(\exists\, \mathbf{v})\mathbf{R}$.

REGLA DE LA GENERALIZACIÓN EXISTENCIAL: *EG. Si una fórmula* **S** *resulta de una fórmula* **R** *sustituyendo todas las incidencias en* **R** *de un nombre ambiguo (o propio), por una variable* **v**, *entonces* $(\exists\, \mathbf{v})\mathbf{S}$ *es deducible de* **R**.

En el siguiente silogismo se da otro ejemplo de aplicación de estas dos reglas.

EJEMPLO 5. *Todos los mamíferos son animales. Algunos mamíferos son bípedos. En consecuencia, algunos animales son bípedos.*

{1}	(1) $(x)(Mx \rightarrow Ax)$	P
{2}	(2) $(\exists x)(Mx \,\&\, Tx)$	P
{2}	(3) $M\alpha \,\&\, T\alpha$	2 *ES*
{1}	(4) $M\alpha \rightarrow A\alpha$	1 *US*
{1, 2}	(5) $A\alpha \,\&\, T\alpha$	3, 4 T
{1, 2}	(6) $(\exists x)(Ax \,\&\, Tx)$	5 *EG*

Continuando las observaciones que se acaban de hacer, nótese que

es necesario que la línea (3) venga antes de la (4). Si se asienta primero la línea (4), entonces, de acuerdo con *ES*, se necesitaría otro nombre ambiguo para la línea (3). La estrategia adecuada consiste en, cada vez que sea posible, suprimir los cuantificadores existenciales *antes* de los cuantificadores universales.

El siguiente ejemplo ilustra cómo se manejan los cuantificadores de una fórmula.

EJEMPLO 6. *Todos los secuaces de Aristóteles estiman a los secuaces del Aquinatense. Ninguno de los secuaces de Aristóteles estima a algún filósofo idealista. Además, Aristóteles tiene secuaces. En consecuencia, ninguno de los secuaces del Aquinatense es filósofo idealista.*

{1}	(1) $(x)(Ax \rightarrow (y)(Qy \rightarrow Lxy))$	P
{2}	(2) $(x)(Ax \rightarrow (z)(Iz \rightarrow -Lxz))$	P
{3}	(3) $(\exists x)(Ax)$	P
{3}	(4) $A\alpha$	3 *ES*
{1}	(5) $A\alpha \rightarrow (y)(Qy \rightarrow L\alpha y)$	1 *US*
{2}	(6) $A\alpha \rightarrow (z)(Iz \rightarrow -L\alpha z)$	2 *US*
{1, 2}	(7) $(y)(Qy \rightarrow L\alpha y)$	4, 5 T
{2, 3}	(8) $(z)(Iz \rightarrow -L\alpha z)$	4, 6 T
{1, 3}	(9) $Qy \rightarrow L\alpha y$	7 *US*
{2, 3}	(10) $Iy \rightarrow -L\alpha y$	8 *US*
{1, 2, 3}	(11) $Qy \rightarrow -Iy$	9, 10 T
{1, 2, 3}	(12) $(y)(Qy \rightarrow -Iy)$	11 *UG*

La línea (2) deliberadamente se ha escrito usando '*z*' en vez de '*y*', a fin de suministrar otra ilustración de cómo se usa la especificación universal *US* para cambiar variables. La observación más importante que debe hacerse con relación a este último ejemplo es que un cuantificador puede ser suprimido sólo cuando está al comienzo de una fórmula y su alcance comprende toda la fórmula. Es un serio error aplicar la especificación universal o la existencial *US* o *ES* cuando no se satisface esta condición. Así, en el Ej. 6, sería un error aplicar la especificación universal *US* a la línea (5) a fin de suprimir el cuantificador '(y)', pues este cuantificador no está al principio de la fórmula. Se habría cometido un error semejante aplicando la especificación universal *US* a la línea (6), a fin de eliminar el cuantificador '(z)'. En el contexto del presente ejemplo esa aplicación de la especificación universal no hubiera dado como consecuencia una conclusión falsa, pero en otros contextos sí, y esto es por lo que debe evitarse.[21]

[21] Véanse los Ejercicios 6 y 7 del § 4.6.

Es necesario hacer notar que aunque la regla de la generalización existencial se usa en primer lugar para eliminar nombres ambiguos y dar paso a cuantificadores existenciales, ocasionalmente resulta necesario poder aplicar un cuantificador existencial a una variable libre. La justificación intuitiva de esta última aplicación radica en que una aserción sobre una cosa arbitraria implica una aserción sobre *alguna* cosa. Formalmente,* esa aplicación de la generalización existencial *EG* se justifica si se considera un nombre ambiguo que no ocurre para nada en la fórmula en cuestión. Por ejemplo, '$(\exists x)Fx$' es derivable de 'Fx' por generalización existencial *EG*. Sustituimos todas las incidencias de 'α' por 'x', digamos, en 'Fx'. Puesto que no hay incidencia de 'α' en 'Fx', la sustitución termina donde se comenzó. En el Cap. 5, en el apartado sobre teoremas de lógica, se dan algunas aplicaciones de este gambito.

Si interpretamos los nombres ambiguos en la misma forma en que los nombres propios y las variables libres, entonces no todas las líneas de una deducción son una consecuencia lógica de la conjunción de las premisas de que dependen. Por ejemplo, de la premisa '$(\exists x) Fx$' derivamos por especificación existencial '$F\alpha$', pero obviamente si 'α' se interpreta justamente como un nombre propio, entonces '$F\alpha$' no es una consecuencia lógica de '$(\exists x)Fx$'. No obstante, esta interpretación es la más natural y el procedimiento más simple consiste en debilitar la exigencia de que todas las líneas de una derivación sean una consecuencia lógica de las premisas de que dependen. Lo que podemos probar es que si una fórmula de una deducción no contiene nombres ambiguos y tampoco los contienen sus premisas, entonces es una consecuencia lógica de las premisas. Este estado de cosas es de hecho intuitivamente satisfactorio, pues en cualquier argumento válido relativo a principios teóricos de una disciplina comenzamos con premisas y terminamos con una conclusión que no contiene nombres ambiguos. La medula de la cuestión en cuanto al uso de nombres ambiguos es que suministran un método fácil de inferir conclusiones y la prueba real de su legitimidad consiste en la validez de la conclusión obtenida, no en el estado de las líneas intermedias.

Por otra parte, el estado lógico de la relación entre premisas y conclusión, una de las cuales cuando menos contiene un nombre ambiguo, puede caracterizarse en forma un tanto sencilla, en función de lo que llamaremos la *oclusión inambigua* de una fórmula. La idea

* El autor hace un uso muy acertado de los vocablos "intuitivo" y "formal", contraponiendo su significado. Por ejemplo, una interpretación intuitiva de una oración sería una interpretación en función de su contenido, esto es, una interpretación "material", epistemológica. Una interpretación formal sería una interpretación en función de las leyes lógicas de validez, con abstracción de su contenido. (N. del T.)

consiste en eliminar los nombres ambiguos por medio de cuantificadores existenciales, pero solamente en un cierto orden relativo a los cuantificadores universales, y determinado por los subíndices de los nombres ambiguos.[22] Por ejemplo, la oclusión inambigua, o para mayor brevedad, la oclusión-*U* de '$H\alpha x$' es '$(\exists y)(x)Hyx$'. La oclusión-*U* de '$H\alpha_x x$' es '$(\exists y)Hyx$'. En términos generales, la oclusión-*U* de una fórmula **S** es la fórmula que se obtiene de **S**, de conformidad con la siguiente receta: si **S** no contiene nombres ambiguos, es su propia oclusión-*U*; si **S** tiene cuando menos un nombre ambiguo, cuantifíquense primero universalmente todas las variables libres de **S** que no aparezcan como subíndices; a continuación reemplácense los nombres ambiguos que lleven más subíndices por una variable apropiada, existencialmente cuantificada (si hay dos o más nombres ambiguos con el número máximo de subíndices, reemplácese el que aparezca primero en **S**); repítase este procedimiento hasta que hayan sido eliminados todos los nombres ambiguos.[23]

Es posible entonces establecer el siguiente resultado respecto de la relación lógica entre las premisas y la conclusión obtenida: la oclusión-*U* de la implicación cuyo antecedente es la conjución de las premisas y cuyo consecuente es la conclusión, es universalmente válida. Dicho de otro modo, toda línea de deducción es una consecuencia de las premisas correspondientes al conjunto de números asentados a su izquierda, en el sentido siguiente: la oclusión-*U* de la implicación consistente en la conjunción de las premisas como antecedente y la línea dada como consecuente, es universalmente válida y por tanto verdadera en toda interpretación correspondiente a un dominio no vacío de individuos.

EJERCICIOS

Constrúyase (de ser posible), una deducción correspondiente a los argumentos siguientes. Si no se sigue una conclusión, dése una interpretación aritmética que pruebe que no se sigue.

1. Alguna gente tonta bebe whiskey. Algunos estudiantes no beben whiskey. En consecuencia, algunos estudiantes no son tontos. (Fx, Wx, Sx)

2. Algunos pugilistas son fuertes. Algunos policías son fuertes. En consecuencia, algunos policías son pugilistas. (Bx, Sx, Px)

3. Algunos temas científicos no son interesantes, pero todos los temas científicos son edificantes. En consecuencia, algunas cosas edificantes no son interesantes. (Sx, Ix, Ex)

[22] La interpretación intuitiva de los subíndices que se usan para indicar relaciones de dependencia, se explica al principio del § 4.5, en la Pág. 126

[23] Es fácil establecer convenciones adecuadas sobre cuál debe ser la ordenación alfabética de las variables para que la oclusión-*U* resulte una fórmula única.

4. Ninguna persona inteligente que bebe con exceso come también con exceso. Algunas personas prudentes comen con exceso. En consecuencia, algunas personas prudentes no son inteligentes. (Ix, Dx, Ex, Px)

5. Ninguna mujer de cabello rojo usa tónico para el cabello, pero algunas mujeres de cabello rojo usan perfume. En consecuencia, algunas mujeres usan perfume, y no usan tónico para el cabello. (Rx, Wx, Hx, Px)

6. Algunos soldados son héroes. Algunos soldados no son valientes. En consecuencia, algunos héroes no son valientes. (Sx, Hx, Bx)

7. Ningún funcionario electo vive en California. Algunas personas que viven en California son posibles candidatos. En consecuencia, algunos candidatos posibles no son funcionarios electos. (Nx, Lx, Cx)

8. Todo miembro del comité político es republicano. Algunos miembros del comité de impuestos no son republicanos. En consecuencia, algunos miembros del comité de impuestos no son miembros del comité político. (Px, Rx, Tx)

9. Algunos miembros del comité de impuestos están en el comité político. Algunos miembros del comité político son demócratas. En consecuencia, algunos miembros del comité de impuestos son demócratas. (Tx, Px, Dx)

10. Todo miembro del comité político es o demócrata o republicano. Algunos miembros de este comité político son ricos. Adams no es demócrata pero es rico. En consecuencia, Si Adams es miembro del comité político, es republicano. (Px, Dx, Rx, Wx, a)

11. Si alguien estima al Aquinatense, entonces no estima a Kant. Todos estiman a Kant o a Russell. Algunos no estiman a Russell. En consecuencia, algunos no estiman al Aquinatense. (Ax, Kx, Rx)

12. Todos los tejanos hablan a cualquiera a quien conocen íntimamente. Ningún tejano habla a alguien que no es habitante del sur. En consecuencia, los tejanos sólo conocen íntimamente a habitantes del sur. $(Tx, Sxy, Kxy\ Ux)$

13. Cualquiera que es culpable de robo no es culpable de estafa. Cada uno de los prisioneros de la rueda de presos es culpable de robo o de estafa. Algunos prisioneros de la rueda de presos son culpables de estafa. En consecuencia, algunos prisioneros de la rueda de presos no son culpables de robo. (Lx, Ox, Px)

14. Algunas personas que compran de buena fe obtienen un título legítimo de propiedad. Ninguna persona que compra por un precio insignificante obtiene un título legítimo de propiedad. En consecuencia, ninguna persona que compra por un precio insignificante compra de buena fe.

15. En la psicología de la percepción así como en epistemología, se discute a menudo la relación entre objetos consistente en la semejanza de color. Por ejemplo, dos tonos de rojo son semejantes; sin embargo, como se concibe comúnmente, la relación de semejanza no es transitiva, esto es, el color a es semejante al b, b es semejante al c, pero a no es semejante al c. De manera que se suscita en forma natural el problema de definir una relación transitiva de *semejanza exacta* o igualdad entre colores. Sea 'L' el predicado que representa la primera relación de semejanza y 'E' el predicado *es igual*. Introducimos entonces la definición

$$(x)(y)(xEy \leftrightarrow (z)(xLz \leftrightarrow yLz) \,\&\, xLy \,\&\, yLx).$$

Usando esta definición como premisa única, pruébese que 'E' es simétrica y

transitiva, esto es, pruébese que:

(a) $(x)(y)(xEy \rightarrow yEx)$
(b) $(x)(y)(z)(xEy \,\&\, yEz \rightarrow xEz)$.[24]

16. Con referencia al ejercicio precedente dése una interpretación de 'L' y 'E' para probar que la oración '$(x)(xEx)$' no es una consecuencia lógica de la equivalencia que define a 'E' en función de 'L'.

§4.4. Intercambio de cuantificadores. Del estudio del significado usual, resulta claro que la oración:

(1) Todo matemático admira a Arquímedes.

es equivalente a:

(2) No es cierto que haya un matemático que no admire a Arquímedes. No obstante, (1) lleva un cuantificador universal y (2) un cuantificador existencial. La representación simbólica de (1) es:

(3) $(x)(Mx \rightarrow Ax),$

y la representación simbólica de (2) es

(4) $-(\exists x)(Mx \,\&\, -Ax).$

Además, puesto que 'Mx & $-Ax$' es tautológicamente equivalente a '$-(Mx \rightarrow Ax)$'*, (4) es equivalente a:

(5) $-(\exists x)-(Mx \rightarrow Ax).$

La estrecha relación entre (3) y (5) sugiere la regla:

> Q1. *Si* **v** *es cualquier variable y si la fórmula* **S** *resulta de* **R** *reemplazando cuando menos una incidencia del cuantificador universal* (**v**) *por* -($\exists$ **v**)-, *entonces* **S** *es derivable de* **R**, *y recíprocamente.*

En el ejemplo anterior v es, por supuesto, la variable 'x'. Usando los métodos del § 5.3 posterior, podemos deducir la regla Q1 de las re-

[24] Este ejercicio fue sugerido por Leo Simons. Para una aplicación de estas nociones en epistemología, véase Hilary Putnam. "Reds, Greens and Logical Analysis", *Philosophical Review*, Vol. 65 (1956). Págs. 206-217.

* En vista del frecuente e importante uso de la tautología $(Mx \,\&\, -A) \leftrightarrow -(Mx \rightarrow Ax)$ nos ha parecido pertinente dar su correspondinte tabla de verdad.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mx	Ax	-Ax	Mx & -Ax	Mx → Ax	-(Mx → Ax)	(4)↔(6)
T	T	F	F	T	F	T
T	F	T	T	F	T	T
F	T	F	F	T	F	T
F	F	T	F	T	F	T

(N. del T.)

glas dadas anteriormente, pero para los fines presentes podemos aceptarlo como suficientemente bien fundamentado para necesitar una ulterior justificación. Aseverar algo de todas las x es exactamente lo mismo que negar que haya una x que no satisfaga esta aseveración.

Por mèdio del siguiente ejemplo se sugiere una regla similar. Está claro que la oración

(6) Hay un becerro de dos cabezas

es equivalente a:

(7) No es cierto que no haya ningún becerro de dos cabezas

Podemos simbolizar (6) por:

(8) $(\exists x)(Cx \ \& \ Tx)$,

y (7) por:

(9) $-(x)(Cx \rightarrow -Tx)$.

Puesto que '$Cx \rightarrow -Tx$' es tautológicamente equivalente a '$-(Cx \ \& \ Tx)$', (9) es equivalente a:

(10) $-(x)-(Cx \ \& \ Tx)$.

La relación entre (8) y (10) sugiere la regla:

> Q2. *Si* **v** *es cualquier variable y si la fórmula* **S** *resulta de* **R**, *reemplazando cuando menos una incidencia del cuantificador existencial* (∃ **v**) *por* -(**v**)-, *entonces* **S** *es derivable de* **R**, *y recíprocamente.*

El significado de Q2 es el de que aseverar que hay una x que tiene cierta propiedad equivale a aseverar que no toda x está sin esta propiedad.

En la discusión de Q1 y de Q2 hemos hecho uso de ciertas equivalencias tautológicas en una forma intuitivamente inobjetable. Una regla de inferencia que justifica este uso es la siguiente.

> Regla de las equivalencias tautológicas. T. E. *Si la fórmula* **P** *interviene como parte de la fórmula* **R**, *si la fórmula* **Q** *es tautológicamente equivalente a* **P**, *y si la fórmula* **S** *resulta de* **R** *reemplazando cuando menos una incidencia de* **P** *en* **R** *por* **Q**, *entonces* **S** *es deducible de* **R**, *y recíprocamente.*

Así pues, Q1 y T. E. justifican formalmente la inferencia de (4) de (3); y Q2 y T. E. apoyan juntamente la inferencia de (9) de (8). Las tres reglas introducidas se usan en la siguiente inferencia.

EJEMPLO 7. *Si hay un tribunal federal que confirme la sentencia, entonces están equivocados todos los miembros de la barra. Sin embargo, algunos miembros de la barra no están equivocados. En consecuencia, ningún tribunal federal confirmará la sentencia.*

{1}	(1) $(\exists x)(Fx \mathbin{\&} Sx) \rightarrow (y)(My \rightarrow Wy)$	P
{2}	(2) $(\exists y)(My \mathbin{\&} -Wy)$	P
{2}	(3) $-(y)-(My \mathbin{\&} -Wy)$	2 Q2
{2}	(4) $-(y)(My \rightarrow Wy)$	3 T.E.
{1, 2}	(5) $-(\exists x)(Fx \mathbin{\&} Sx)$	1, 4 T
{1, 2}	(6) $--(x)-(Fx \mathbin{\&} Sx)$	5 Q2
{1, 2}	(7) $(x)-(Fx \mathbin{\&} Sx)$	6 T
{1, 2}	(8) $(x)(Fx \rightarrow -Sx)$	7 T.E.

Al deducir la línea (6) de la (5), el cuantificador existencial '$(\exists x)$' fue reemplazado por '$-(x)-$', aunque no estaba al principio de la fórmula en (5). Ese intercambio contextual es permitido por Q1 y Q2. Esto diferencia estas reglas de *US* y *ES* que se aplican sólo cuando el cuantificador está enfrente de toda la fórmula.

En el § 5.3 se enuncian algunas reglas más.

EJERCICIOS

1. Si todos los miembros de la barra están equivocados, entonces hay un tribunal federal que confirmará la sentencia. Pero ningún tribunal federal confirmará la sentencia. En consecuencia, algunos miembros de la barra no están equivocados. (Mx, Wx, Fx, Sx)

2. Si todos los testigos dicen la verdad, entonces Bluenose será encontrado culpable o se dividirá el jurado. Pero no se dividirá el jurado. En consecuencia, o Bluenose será encontrado culpable, o algún testigo no está diciendo la verdad. (Wx, Tx, B, H)

3. Si Bluenose es culpable, entonces ningún testigo miente, a menos que tenga miedo. Hay un testigo que tiene miedo. En consecuencia, Bluenose no es culpable.

§4.5. Inferencias generales. Limitando la clase de inferencias consideradas, pudimos limitar a dos el número de restricciones impuestas a las reglas que rigen a los cuantificadores: una impuesta a la generalización universal *UG* (la restricción de imprecisión) y la otra a la especificación existencial *ES* (introducir con cada aplicación un nuevo nombre ambiguo). A fin de ampliar nuestras reglas a las inferencias generales y no violar el Criterio I, esto es, no permitir la inferencia de una conclusión falsa, de premisas verdaderas, debemos imponer cinco restricciones más: Una a la especificación universal *US*,

otra a la generalización universal *UG* y tres a la generalización existencial *EG*. También debe complicarse un poco más el uso y el enunciado de la especificación existencial *ES*.

Estas cinco restricciones sólo implican una innovación en las notaciones. Introducimos el uso de variables como *subíndices* para evitar que *ES* nos lleve a una contradicción. Un ejemplo sencillo ilustrará la dificultad y nuestro modo particular de resolverla. Es una verdad aritmética que no existe un número máximo, esto es,

(1) $$(x)(\exists y)(x < y).$$

Aplicamos *US* y obtenemos:

(2) $$(\exists y)(x < y).$$

Si aplicamos *ES* como se define en §4.3, obtenemos a continuación:

(3) $$x < \alpha,$$

y entonces por *EG* obtenemos la falsa aserción de que

(4) $$(\exists x)(x < x).$$

Cerramos el paso a esta inferencia falaz requiriendo que si se introduce por *ES* un nombre ambiguo, incluya como subíndices todas las variables libres existentes en la fórmula original. Así, en (2), 'x' es una variable libre. De aquí que ahora no obtengamos (3), sino

(3') $$x < \alpha_x.$$

Finalmente, restringimos el uso de *EG*.

> PRIMERA NUEVA RESTRICCIÓN IMPUESTA A EG. *No se puede aplicar a una fórmula dada un cuantificador existencial que use una variable que aparece como subíndice en la fórmula.*

De aquí que no podamos inferir (4) de (3'), puesto que 'x' aparece como subíndice en (3').

Debe hacerse hincapié en que este uso de subíndices no es un artificio arbitrario, sino que se aproxima mucho a los métodos matemáticos. En (3), la determinación de α depende del número x. En (3') hacemos explícita esta dependencia de α, escribiendo 'α_x'. Puesto que α_x depende específicamente de x, la variable 'x' no puede ser cuantificada mientras estén presentes nombres ambiguos en los cuales aparezca esa variable como subíndice.

Repetimos en forma suscinta el procedimiento para la aplicación de subíndices.

Cuando se introduce por ES un nombre ambiguo, debe incluir como subíndices todas las variables libres que aparecen en la fórmula a la cual se aplica ES.[25]

Debe hacerse notar que si tenemos una línea como (3') podemos aplicar un cuantificador existencial usando cualquier variable distinta de 'x'. Por ejemplo, es completamente correcto deducir de (3'):

$$(\exists z)(x < z).$$

La restricción impuesta a *EG* debe extenderse también a *UG*, como lo demuestra la siguiente línea de inferencia. Volvamos a (1), que asevera que no existe ningún número máximo absoluto. Como anteriormente, deducimos (2) por *US*:

$$(\exists y)(x < y),$$

y aplicamos ahora la regla nuevamente modificada de especificación existencial, para obtener (3'):

$$x < \alpha_x.$$

Hasta ahora todo va bien, pero aplicamos la generalización universal:

(5) $$(x)(x < \alpha_x),$$

y a este resultado le aplicamos *EG*, a fin de obtener la oración falsa:

(6) $$(\exists y)(x)(x < y).$$

Se cierra el paso a este argumento falaz prohibiendo el paso de (3') a (5). Esto es, restringimos *UG* en la misma forma que *EG*.

NUEVA RESTRICCIÓN IMPUESTA A *UG*. *No podemos aplicar a una fórmula dada un cuantificador universal que use una variable que aparezca como subíndice en la fórmula.*

Este último ejemplo que justifica la restricción impuesta a *UG*, puede formularse como una deducción falaz para poner de manifiesto la manera como funciona la nueva restricción, en una deducción.

EJEMPLO FALAZ DE RESTRICCIÓN MEDIANTE SUBÍNDICE IMPUESTA A *UG*.

{1}	(1) $(x)(\exists y)(x < y)$	P
{1}	(2) $(\exists y)(x < y)$	1 *US*
{1}	(3) $x < \alpha_x$	2 *ES*

[25] Debe hacerse la observación de que cualquier incidencia de una variable como subíndice es una incidencia libre. Así, de '$(\exists y)(\alpha_x < y)$' deducimos '$\alpha_x < \beta_x$', y no '$\alpha_x < \beta$'.

{1}	(4) $(x)(x < \alpha_x)$	3 *UG* (falazmente)
{1}	(5) $(\exists y)(x)(x < y)$	4 *EG*

La variable '*x*' ocurre como subíndice en la línea (3) y es, en consecuencia, incorrecto aplicar *UG* para deducir (4).

La justificación de la nueva restricción impuesta a la especificación universal es fácil de ilustrar. Comenzamos de nuevo con la premisa verdadera de que no hay un número máximo absoluto:

$$(x)(\exists y)(x < y),$$

aplicamos *US*, y reemplazamos '*x*' por '*y*', para obtener la oración falsa:

$$(\exists y)(y < y).$$

La deducción falaz es como sigue:

{1}	(1) $(x)(\exists y)(x < y)$	P
{1}	(2) $(\exists y)(y < y)$	1 *US* (falazmente)

La clase de restricción que es necesario imponer a *US* resulta obvia de este ejemplo: no se use como sustituto una variable que quede ligada.* Así, en la línea (2) anterior '*y*' reemplaza a '*x*' y queda ligada, lo que constituye la fuente de la inferencia inválida.

> NUEVA RESTRICCIÓN IMPUESTA A *US*. *No se use como sustituto un término que contenga una variable que, como resultado de la sustitución, quede ligada por un cuantificador de la fórmula original.*

La restricción se enuncia para términos, pues no solamente la sustitución de una variable podría acarrear dificultades, sino también la sustitución de un término que no sea una variable. Por ejemplo, también habría sido falaz en el ejemplo anterior sustituir a '*x*' por '*x* + *y*', pues '*y*' queda ligada en (2), ya por sí misma, ya como parte de un término. Habríamos obtenido:

$$(\exists y)(x + y < y).$$

Generalizando universalmente, tenemos:

$$(x)(\exists y)(x + y < y),$$

---

* Esto resulta más claro si se da la forma siguiente al ejemplo ofrecido por el autor:

{1}	$(x)(\exists y)(x < y)$	P
{2}	$(\exists y)(x < y)$	US
{3}	$(\exists y)(y < y)$	Sustitución

La "fórmula original" a que se refiere la regla de la restricción es el renglón (2), que, por lo demás, es la única fórmula en el anterior desarrollo, pues los renglones (1) y (3) propiamente son oraciones. (N. del T.)

y por especificación universal deducimos a continuación la oración falsa:

$$(\exists y)(0 + y < y).$$

Nuestra siguiente restricción, la siguiente restricción nueva de la generalización existencial, es idéntica a la nueva restricción de la especificación universal: No se use como variable sustituta una variable que sea capturada por un cuantificador ya presente en la fórmula. El siguiente ejemplo muestra la necesidad de imponer esta restricción a *EG*. Comenzamos con la oración verdadera de aritmética:

(1) $$(\exists x)(y)(x + y = y).$$

Para cerciorarse de que (1) es verdadera, tómese '*x*' como '0'. Aplicamos especificación existencial para obtener:

(2) $$(y)(\alpha + y = y).$$

Aplicamos *EG* a (2) y nos metemos en dificultades:

(3) $$(\exists y)(y)(y + y = y).$$

Finalmente, aplicamos *ES* a (3) y obtenemos la oración falsa:

(4) $$(y)(y + y = y).$$

Puesto que en la línea (2) 'α' cae dentro del alcance del cuantificador '(y)', es un error reemplazar 'α' por 'y' y añadir el cuantificador existencial '$(\exists\ y)$' en la línea (3). Nótese que deducimos trivialmente por *ES* la línea (4) de la (3), pues no hay variables libres en '$(y)(y < y)$' que reemplazar por un nombre ambiguo. Las fórmulas como (3) satisfacen la definición de fórmula, pero las reglas de inferencia están estructuradas de tal modo que el cuantificador externo es esencialmente redundante. Cuando la incidencia de una variable cae dentro del alcance de dos cuantificadores, siempre tiene la preferencia el cuantificador interno. Resumiendo, tenemos:

SEGUNDA NUEVA RESTRICCIÓN IMPUESTA A *EG*. *No se reemplace un nombre ambiguo por una variable que resulte ligada por un cuantificador de la fórmula original.*

Nuestra restricción final será la tercera restricción impuesta a *EG*. Se refiere a generalizar existencialmente, en forma simultánea, una variable imprecisa y un nombre ambiguo (o propio). Para que se susciten dificultades, el nombre debe de hecho aparecer cuando menos una vez en la fórmula que se generaliza. Un ejemplo aritmético que usa solamente el predicado '*O*' y en que '*Ox*' significa que *x* es un entero impar, ilustrará la dificultad.

{1}	(1) $(\exists x)-Ox$	P
{2}	(2) Ox	xP
{1}	(3) $-O\alpha$	1 *ES*
{1, 2}	(4) $Ox \& -O\alpha$	x 2, 3 T
{1, 2}	(5) $(\exists x)(Ox \& -Ox)$	4 *EG* (falazmente)

La línea (1) es una verdad de aritmética y si interpretamos la variable libre 'x' de (2) como '1', entonces (2) es verdadera, pero (5) es notoriamente falsa. La línea (4) no debió haber sido generalizada existencialmente en función de 'x' para dar (5), puesto que aparecen juntamente 'α' y 'x' imprecisas. Esta restricción final no es violada a menudo en la práctica.[26] Su enunciado formal es:

TERCERA NUEVA RESTRICCIÓN IMPUESTA A *EG*: *No se use una variable imprecisa de una fórmula para eliminar de la fórmula la incidencia de un nombre.*

No hay peligro en generalizar existencialmente en función de una variable imprecisa presente en una fórmula, cuando esta generalización no elimina alguna incidencia de un nombre. Por ejemplo, de la premisa 'Fx' podemos inferir '$(\exists x)Fx$', aun cuando 'x' está imprecisa en la premisa. Si tenemos el propósito de eliminar 'α' con esta generalización, 'α' no aparece de hecho en 'Fx'. La eliminación no se efectúa, y no hay dificultad.

Concluimos este apartado con varios ejemplos que ilustran el uso de las reglas en su forma final, esto es, con subíndices en los nombres ambiguos cuando sea necesario y con las cinco nuevas restricciones satisfechas.

El lógico inglés del siglo XIX, De Morgan, sostuvo que la lógica silogística clásica era demasiado débil para deducir que todas las cabezas de caballos eran cabezas de animales, de la premisa de que todos los caballos son animales. La deducción deseada ilustra el uso de los subíndices. Usamos 'H' para 'es una cabeza de' y 'P' para 'es un caballo', y a continuación transcribimos como sigue la conclusión deseada:

$$(x)[(\exists y)(Py \& Hxy) \rightarrow (\exists y)(Ay \& Hxy)].$$

EJEMPLO 8. *Todos los caballos son animales. En consecuencia, todas las cabezas de caballo son cabezas de animales.*

{1}	(1) $(x)(Px \rightarrow Ax)$	P
{2}	(2) $(\exists y)(Py \& Hxy)$	xP
{2}	(3) $P\alpha_x \& Hx\alpha_x$	x2 *ES*
{1}	(4) $P\alpha_x \rightarrow A\alpha_x$	1 *US*

[26] Rigurosamente hablando, esta restricción se necesitó en el §4.3.

{1, 2}	(5) $A\alpha_x \,\&\, Hx\alpha_x$	x3, 4 T
{1, 2}	(6) $(\exists y)(Ay \,\&\, Hxy)$	x5 *EG*
{1}	(7) $(\exists y)(Py \,\&\, Hxy) \rightarrow (\exists y)(Ay \,\&\, Hxy)$	2, 6 C.P.
{1}	(8) $(x)[(\exists y)(Py \,\&\, Hxy) \rightarrow (\exists y)(Ay \,\&\, Hxy)]$	7 *UG*

En la línea (2) introducimos el antecedente de la conclusión como una premisa de trabajo. Nótese que 'x' está libre en la premisa. En la línea (3) usamos *ES* para introducir el nombre ambiguo 'α_x'. Intuitivamente, α_x es el caballo del cual x es la cabeza y el subíndice 'x' se usa para indicar la dependencia.

El siguiente ejemplo trata del filósofo nominalista del siglo XIV Guillermo de Ockham y del gran filósofo del siglo XVII Tomás Hobbes.

EJEMPLO 9. *Ninguno de los secuaces de Ockham estima a ningún realista. Cualquiera de los secuaces de Ockham estima cuando menos a uno de los secuaces de Hobbes. Además, Ockham tiene secuaces. En consecuencia, algunos de los secuaces de Hobbes no son realistas.*

{1}	(1) $(x)(Ox \rightarrow (y)(Ry \rightarrow -Lxy))$	P
{2}	(2) $(x)(Ox \rightarrow (\exists y)(Hy \,\&\, Lxy))$	P
{3}	(3) $(\exists x)(Ox)$	P
{3}	(4) $O\alpha$	3 *ES*
{1}	(5) $O\alpha \rightarrow (y)(Ry \rightarrow -L\alpha y)$	1 *US*
{2}	(6) $O\alpha \rightarrow (\exists y)(Hy \,\&\, L\alpha y)$	2 *US*
{2, 3}	(7) $(\exists y)(Hy \,\&\, L\alpha y)$	4, 6 T
{2, 3}	(8) $H\beta \,\&\, L\alpha\beta$	7 *ES*
{1, 3}	(9) $(y)(Ry \rightarrow -L\alpha y)$	4, 5 T
{1, 3}	(10) $R\beta \rightarrow -L\alpha\beta$	9 *US*
{1, 2, 3}	(11) $-R\beta$	8, 10 T
{1, 2, 3}	(12) $H\beta \,\&\, -R\beta$	8, 11 T
{1, 2, 3}	(13) $(\exists x)(Hx \,\&\, -Rx)$	12 *EG*

Aunque en este ejemplo se necesitan dos nombres ambiguos, ninguno de ellos requiere subíndices, puesto que no hay variables libres ni en la línea (3) ni en la (7).

La siguiente sencilla deducción demuestra que no hay dificultad en manejar nombres ambiguos con más de un subíndice. El ejemplo trata de puntos sobre una línea: xPy cuando x precede a y sobre la línea; $B(x, y, z)$ cuando y está entre x y z. Los hechos aseverados en las premisas son obvios, como también lo es la conclusión.

EJEMPLO 10. *Para toda x y para toda y, si x precede a y, no se da el caso de que y preceda a x. Para toda x, toda y y toda z, si x precede a y y y precede a z, entonces x precede a z. Para toda x, y toda y, si x precede a y, entonces x no es igual a y. Para toda x, toda y y toda z, si y está entre x y z, entonces, o x precede a y y precede a z, o z precede a y y precede a x. Para toda x y toda z, si x ≠ z, entonces hay una y tal que y está entre x y z. En consecuencia, para toda x y toda z, si x precede a z, entonces hay una y tal que x precede a y y y precede a z.*

{1}	(1)	$(x)(y)(xPy \rightarrow -yPx)$	P
{2}	(2)	$(x)(y)(z)(xPy \,\&\, yPz \rightarrow xPz)$	P
{3}	(3)	$(x)(y)(xPy \rightarrow x \neq y)$	P
{4}	(4)	$(x)(y)(z)(B(x, y, z) \rightarrow (xPy \,\&\, yPz) \vee (zPy \,\&\, yPx))$	P
{5}	(5)	$(x)(z)(x \neq z \rightarrow (\exists y)B(x, y, z))$	P
{6}	(6)	xPz	x, z P
{3}	(7)	$xPz \rightarrow x \neq z$	3 *US* $x/x, z/y$
{5}	(8)	$x \neq z \rightarrow (\exists y)B(x, y, z)$	5 *US* $x/x, z/z$
{3, 5, 6}	(9)	$(\exists y)B(x, y, z)$	x, z 6, 7, 8 T
{3, 5, 6}	(10)	$B(x, \alpha_{xz}, z)$	x, z 9 *ES*
{4}	(11)	$B(x, \alpha_{xz}, z) \rightarrow (xP\alpha_{xz} \,\&\, \alpha_{xz}Pz) \vee (zP\alpha_{xz} \,\&\, \alpha_{xz}Px)$	4 *US* $x/x, \alpha_{xz}/y, z/z$
{3, 4, 5, 6}	(12)	$(xP\alpha_{xz} \,\&\, \alpha_{xz}Pz) \vee (zP\alpha_{xz} \,\&\, \alpha_{xz}Px)$	x, z 10, 11 T
{1}	(13)	$xPz \rightarrow -zPx$	1 *US* $x/x, z/y$
{1, 6}	(14)	$-zPx$	x, z 6, 13 T
{2}	(15)	$zP\alpha_{xz} \,\&\, \alpha_{xz}Px \rightarrow zPx$	2 *US* $z/x, \alpha_{xz}/y, x/z$
{1, 2, 6}	(16)	$-(zP\alpha_{xz} \,\&\, \alpha_{xz}Px)$	x, z 14, 15 T
{1, 2, 3, 4, 5, 6}	(17)	$xP\alpha_{xz} \,\&\, \alpha_{xz}Pz$	x, z 12, 16 T
{1, 2, 3, 4, 5, 6}	(18)	$(\exists y)(xPy \,\&\, yPz)$	x, z 17 *EG*
{1, 2, 3, 4, 5}	(19)	$xPz \rightarrow (\exists y)(xPy \,\&\, yPz)$	6, 18 C.P.
{1, 2, 3, 4, 5}	(20)	$(x)(z)(xPz \rightarrow (\exists y)(xPy \,\&\, yPz))$	19 *UG*

puesto que 'x' y 'z' están ambas libres en la línea (9) aparecen como subíndices en la línea (10) y subsecuentemente. En las líneas (7), (8), (11), (13) y (15) se introduce cierta nueva rotación a la derecha, para indicar un uso múltiple de *US*. Por ejemplo, al deducir

(7) de (3), se suprimen ambos cuantificadores universales de (3). De acuerdo con una interpretación estricta de *US* no se permite esa doble aplicación, pero está claro que puede fácilmente establecerse una regla derivada que permita esas inferencias. Las nuevas notaciones usadas a la derecha se explican por sí mismas. En la línea (11), por ejemplo, 'x/x, α_{xz}/y, z/z' indica que al aplicar *US* a la línea (4), 'x' y 'z' se sustituyeron trivialmente por sí mismas y 'α_{xz}' se colocó en lugar de 'y'. Cada vez que se hace esa aplicación múltiple de *US* la sustitución se indica a la derecha.

Puesto que la línea (6) es una premisa, las variables 'x' y 'z' están imprecisas en esta línea y en toda línea subsecuente que dependa de esta premisa. El hecho de que en la línea (10) aparezcan estas premisas como imprecisas y como subíndices, es un accidente lógico. No hay una conexión sistemática directa entre ambas operaciones, a no ser que las dos se apliquen a las variables libres.

Cuando las deducciones llegan a tener tantas líneas como el Ej. 10, un plan de trabajo útil es el de hacer el borrador de una línea de inferencia que parezca conveniente, antes de intentar la formulación de una derivación formal detallada. En los Caps. 5 y 7 se discuten *in extenso* formas de acortar o eliminar los pasos de rutina.

EJERCICIOS

Constrúyase (de ser posible), una deducción correspondiente a los argumentos de los Ejercicios 1 a 8. Si el argumento no es válido, pruebe por el método de interpretación.

1. Algunos de los secuaces de Aristóteles estiman a todos los secuaces del Aquinatense. Ninguno de los secuaces de Aristóteles estima a ningún idealista. En consecuencia, ninguno de los secuaces del Aquinatense es idealista. (Ax, Lxy, Qx, Ix)

2. Si gana el equipo, entonces alguno en el terreno trasero es un buen zaguero. Adams es un buen zaguero. En consecuencia, si Adams está en el terreno trasero, gana el equipo. (W, Bx, Tx, a)

3. Ninguna de las pinturas es valiosa, excepto las de combates. Todas las pinturas de combates están pintadas al óleo. Algunas de las pinturas no están pintadas al óleo. Algunas pinturas no tienen marcos. En consecuencia, no es valiosa ninguna de las pinturas no pintadas al óleo. (Px, Vx, Bx, Ox, Fx)

4. Algunos psicólogos admiran a Freud. Algunos psicólogos no estiman a nadie que admire a Freud. En consecuencia, algunos psicólogos no son estimados por todos los psicólogos. (Px, Fx, Lxy)

5. Si Round Robin ganó la carrera, entonces se sintieron felices algunas personas que estaban en el hipódromo. Si todas las personas que apostaron perdieron su dinero, entonces no se sintió feliz ninguna de las personas que estuvieron en el hipódromo. En consecuencia, si Round Robin ganó la carrera, entonces alguien que apostó no perdió su dinero. (R, Tx, Hx, Bx, Lx)

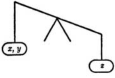
6. Kilroy estuvo aquí. En consecuencia, alguien estuvo aquí. (Hx, K)

7. Todos los buenos críticos estiman a todos los poetas mencionados en la conferencia. Ningún crítico bueno estima a Edgar Guest, aunque Edgar Guest es un poeta. En consecuencia, Edgar Guest no fue mencionado en la conferencia. (Cx, Lxy, Mx, Px, g)

8. Todo empirista filosófico admira a Hume. Algunos idealistas filosóficos no estiman a nadie que admire a Hume. En consecuencia, algunos idealistas filosóficos no estiman a ningún empirista filosófico. (Ex, Hx, Ix, Lxy)

9. La teoría de las mediciones empíricas suministra algunos de los ejemplos más sencillos de teorías científicas no triviales. Este ejercicio trata de una serie de axiomas relativos a la medición de la masa. Intuitivamente, el dominio de individuos es un gran conjunto de objetos físicos que podemos colocar en uno de los dos platillos de una balanza de brazos iguales. Como en el caso de la teoría de la preferencia, 'Q' es un predicado de dos posiciones con propiedades como '$\leq$'. Si entre los dos objetos existe la relación Q, esto es, si xQy, entonces la balanza está en equilibrio horizontal (los objetos tienen la misma masa), o el platillo en que está x está más alto que el platillo en que está y (x tiene menos masa que y). De este modo, mediante una observación cualitativa, podemos siempre decidir si dos objetos físicos están en una relación Q, cuando cada uno de ellos es colocado en el platillo de una balanza. También tenemos una operación $*$ para combinar objetos, a fin de formar nuevos objetos. Así, si $x * yQz$, entonces x y y están en un platillo de la balanza, z en el otro; y el experimentador observa:

o



esto es, la combinación de x y y tiene la misma o menos masa que z. Desde un punto de vista lógico '$*$' es un símbolo de operación que se usa para formar términos como '$x * y$', '$z * x$'. Es importante advertir que '$(x * y) * z$' no es lo mismo que '$x * (y * z)$'; en otras palabras, es lógicamente posible que el orden de agrupación o combinación, como lo indican los paréntesis, pueda significar una diferencia de masa. (Es una consecuencia de los axiomas que no lo signifique, pero tiene que probarse).

Deseamos enunciar cinco axiomas que podríamos esperar que queden satisfechos y a continuación enunciar ocho teoremas que pueden ser deducidos de estos axiomas por las reglas de inferencia que hemos desarrollado. Se necesita un teorema adicional (llamado axioma de Arquímedes) para continuar la lista de teoremas y probar, finalmente, que podemos introducir con propiedad números para medir la masa. No damos aquí este último axioma a causa de su carácter más avanzado.[27]

Los cinco axiomas son:

(1) $(x)(y)(z)(xQy \;\&\; yQz \rightarrow xQz)$

(2) $(x)(y)(z)[(x * y) * z\, Q\, x * (y * z)]$

[27] El conjunto completo de axiomas se encuentra en mi artículo "A Set of Independent Axioms for Extensive Quantities", *Portugaliae Mathemática*, Vol. 10 (1951). Págs. 163-172.

(3) $(x)(y)(z)(xQy \rightarrow x \star z \, Q \, z \star y)$
(4) $(x)(y)(-xQy \rightarrow (\exists z)(x \, Q \, y \star z \,\&\, y \star z \, Q \, x))$
(5) $(x)(y)(-x \star y \, Q \, x)$

El primer axioma dice que la relación de igual, o menor que, en masa es transitiva. El segundo axioma permite reordenaciones de objetos. El tercer axioma asevera que si x es menor o igual que y en masa, entonces no cambiará la igualdad si se combina x con z y z con y. El cuarto axioma dice que si x es estrictamente más pesado que y, entonces puede encontrarse un objeto z que combinado con y equilibre exactamente a x. El quinto axioma dice que todo objeto tiene masa positiva; es decir, que si se combina cualquier objeto x con cualquier objeto y, la combinación es más pesada que x. Tratando los cinco axiomas como premisas, derívense formalmente los ocho teoremas como conclusiones.

(Para evitar una larga deducción, se sugiere que en la deducción de todos los teoremas, menos el primero, se inserte un teorema previo cuando se necesite, justificándolo a la derecha por el número del teorema y listando a la izquierda los teoremas de que depende. Por ejemplo, se usará el Teorema 1 en la prueba del Teorema 2. En el siguiente capítulo se discute sistemáticamente la inserción en una deducción de esos resultados previamente establecidos.)

TEOREMA 1. $(x)(xQx)$
TEOREMA 2. $(x)(y)(x \star y \, Q \, y \star x)$
TEOREMA 3. $(x)(y)(u)(v)(xQy \,\&\, uQv \rightarrow x \star u \, Q \, y \star v)$
TEOREMA 4. $(x)(y)(z)[x \star (y \star z) \, Q \, (x \star y) \star z]$

(Esta derivación es larga)

TEOREMA 5. $(x)(y)(xQy \vee yQx)$

[Sugestión: Use una prueba indirecta y el axioma (4).]

TEOREMA 6. $(x)(y)(z)(x \star z \, Q \, y \star z \rightarrow xQy)$
TEOREMA 7. $(x)(y)(z)(u)(y \star z \, Q \, u \,\&\, xQy \rightarrow x \star z \, Q \, u)$
TEOREMA 8. $(x)(y)(z)(u)(u \, Q \, x \star z \,\&\, xQy \rightarrow u \, Q \, y \star z)$

10. Considérense los cinco axiomas dados en el ejercicio anterior. La siguiente interpretación aritmética prueba que el Axioma (1) es independiente de los cuatro restantes:

(i) El dominio de individuos es el conjunto de enteros positivos.
(ii) $xQy \leftrightarrow x \leq y + 1$.
(iii) $x \star y = x + y + 2$.

Se comprueba fácilmente que el Axioma (1) es falso en esta interpretación y que son verdaderos los cuatro axiomas siguientes. Encuéntrense interpretaciones que prueben que los otros cuatro axiomas son también independientes cada uno. Como sugestión para construir una interpretación adecuada, se da un dominio satisfactorio de individuos para cada axioma:

Para el Axioma (2): El conjunto de todos los números positivos racionales (un número racional es la razón de dos enteros; es un poco difícil encontrar la interpretación correcta de este axioma).

Para el Axioma (3): El conjunto de todos los números positivos racionales.

Para el Axioma (4): El conjunto de todos los enteros positivos, con excepción de uno.

Para el Axioma (5): El conjunto que consiste exclusivamente en el número uno.

§4.6. Resumen de las reglas de inferencia. El propósito de este apartado es el de resumir en forma tabular las siete reglas básicas de inferencia que han sido introducidas en el Cap. 2 y en el presente.

Mejor que repetir la compleja fraseología de sustitución usada en el enunciado inicial de las reglas, es empleada una conveniente notación abreviada. Si **S(v)** es cualquier fórmula en que la variable **v** está libre, entonces **S(t)** es la fórmula que resulta de **S(v)** sustituyendo toda ocurrencia libre de **v** en **S** por el término **t**. Así, si **v** es '*x*', **S(v)** es '*Fx*' y **t** es '*y*', entonces **S(t)** es '*Fy*'. Como segundo ejemplo, si

$$\mathbf{v} = \text{'}z\text{'}$$
$$\mathbf{S(v)} = \text{'}(\exists z)(x + z = 0) \;\&\; y + z > z + 2\text{'}$$
$$\mathbf{t} = \text{'}x + y\text{'}$$

entonces

$$\mathbf{S(t)} = \text{'}(\exists z)(x + z = 0) \;\&\; y + (x + y) > (x + y) + 2\text{'}.$$

De modo semejante, si **S(ν)** es cualquier forma en que aparece el nombre ambiguo **ν**, entonces se obtiene **S(v)** sustituyendo en **S(ν)** si y sólo si toda incidencia de **ν** en **S(ν)** es reemplazada por la variable **v**. Así, si

$$\boldsymbol{\nu} = \text{'}\alpha\text{'}$$
$$\mathbf{S}(\boldsymbol{\nu}) = \text{'}\alpha + x = \beta \;\&\; \alpha > 0\text{'}$$
$$\mathbf{v} = \text{'}x\text{'}$$

entonces

$$\mathbf{S(v)} = \text{'}x + x = \beta \;\&\; x > 0\text{'}.$$

REGLAS BASICAS DE INFERENCIA

ABREVIATURA	REGLA	RESTRICCIÓN
P	Introducción de premisas	Ninguna
T	Uso de tautologías	Ninguna
C.P.	Prueba condicional	Ninguna
US	Especificación universal: de (v)S derivar S(t)	Ninguna incidencia libre de v dentro del alcance de un

		cuantificador que use una variable de t.
UG	Generalización universal: de S derivar (v)S	(1) v no imprecisa (2) v no como subíndice
ES	Especificación existencial: de (∃v)S(v) derivar S(ν)	Que el nombre ambiguo ν no haya sido previamente usado
EG	Generalización existencial: de S(ν) derivar (∃v)S(v)	(1) v no aparece como subíndice (2) Ninguna incidencia del nombre ν queda dentro del alcance de un cuantificador que usa v (3) v no está imprecisa si ν no aparece en S(ν)

IMPRECISIÓN: Una variable libre de una premisa es imprecisa y continúa imprecisa en cualquier línea en que esté libre y que dependa de la premisa.

SUBÍNDICES: Cuando se introduce un nombre ambiguo por especificación existencial, debe tener como subíndices todas las variables libres que aparecen en la fórmula a la cual se aplica la especificación existencial.

EJERCICIOS

¿Qué error se comete y en qué punto en las deducciones falaces 1 a 5?

1. Si se da una cosa cualquiera, existe otra cosa distinta de la primera. En consecuencia, existe algo que es distinto de sí mismo.

{1}	(1) $(x)(\exists y)(x \neq y)$
{1}	(2) $(\exists y)(y \neq y)$

2. Alguien tiene cabello rojo. Alguien tiene cabello castaño. En consecuencia, alguien tiene cabello rojo y cabello castaño.

{1}	(1) $(\exists x)Rx$
{2}	(2) $(\exists x)Bx$
{1}	(3) $R\alpha$
{2}	(4) $B\alpha$
{1, 2}	(5) $R\alpha \,\&\, B\alpha$
{1, 2}	(6) $(\exists x)(Rx \,\&\, Bx)$

3. Algunos hombres tienen más de 1.80 m de alto. En consecuencia, todos los hombres tienen más de 1.80 m de alto.

{1}	(1) $(\exists x)Ox$
{1}	(2) $O\alpha$
{1}	(3) $(x)Ox$

4. Todo hombre es tan alto como sí mismo. En consecuencia, hay un hombre que es tan alto como todos los demás hombres.

{1}	(1) $(x)Txx$
{1}	(2) Tax
{1}	(3) $(x)Tax$
{1}	(4) $(\exists y)(x)Tyx$

5. Todo es idéntico consigo mismo. En consecuencia, todo es idéntico con todo lo demás.

{1}	(1) $(x)(x = x)$
{1}	(2) $x = y$
{1}	(3) $(y)(x = y)$
{1}	(4) $(x)(y)(x = y)$

6. Constrúyase una deducción falaz que viole la restricción de que un cuantificador universal puede ser suprimido por *US* sólo cuando esté al frente de la fórmula y su alcance sea toda la fórmula.

7. Constrúyase una deducción falaz que viole la restricción de que un cuantificador existencial puede ser suprimido por *ES* sólo cuando el cuantificador está al frente de la fórmula y su alcance abarca a toda fórmula.

CAPITULO 5

OTRAS REGLAS DE INFERENCIA

§ 5.1. **Lógica de la identidad.** En el lenguaje cotidiano a menudo colocamos una forma del verbo 'ser' entre dos términos, a fin de indicar que designan la misma entidad. Escribimos, por ejemplo:

(1) Isabel II es la actual reina de Inglaterra,

para significar que Isabel II es la misma persona que la actual reina de Inglaterra. O (usando aquí el tiempo pasado), escribimos:

(2) James Madison fue el cuarto presidente de los Estados Unidos,

para significar que James Madison y el cuarto presidente de los Estados Unidos son la misma persona.

El lenguaje cotidiano da al verbo 'ser' otras muchas aplicaciones. Por ejemplo, algunas veces lo usamos para indicar que cierta entidad tiene cierta propiedad. Por ejemplo, cuando escribimos:

(3) La sal es blanca,

ciertamente no queremos decir que sal es lo mismo que blanco —lo que carecería de sentido—, ni tampoco que sal es lo mismo que blancura —lo que sería falso—, sino que la sal posee la propiedad de ser blanca.

Aunque en el lenguaje usual nos las arreglamos para usar el verbo 'ser' a pesar de ciertas confusiones, en las discusiones filosóficas la ambigüedad de esta palabra puede ser motivo de oscuridad y de confusiones interminables. De manera que para una expresión más exacta es conveniente repartir los diversos significados del verbo 'ser' entre diversas palabras y símbolos que se introducen especialmente para este propósito. En particular, se usan las palabras 'es idéntico', o el símbolo '=' para el primer significado mencionado.

(En el Cap. 9 introduciremos otras palabras y símbolos para otros significados del verbo 'ser'). En vez de (1) escribimos:

Isabel II es idéntica a* la actual reina de Inglaterra,

o simplemente:

Isabel II = la actual reina de Inglaterra.

Y en vez de (2) escribimos:

James Madison es idéntico* al cuarto presidente de las Estados Unidos,

(4) James Madison = el cuarto presidente de los Estados Unidos.

En cambio, no reemplazamos el 'es' de (3) por 'es idéntico a'* o '='; puesto que el 'es' de (3) tiene, como hicimos notar, un significado completamente diferente del 'es' de (1) o del 'fue' de (2).

Algunas veces leeremos el símbolo '=' también como 'igual a'. Pero debe hacerse hincapié en que algunos usan la expresión 'igual a' en un sentido un tanto más amplio; por ejemplo, en geometría elemental dos segmentos de recta *AB* y *CD* se llaman algunas veces iguales si tienen la misma longitud. En cambio, nosotros nunca nos expresaremos de este modo. Si nosotros escribiéramos:

$$AB = CD,$$

o dijéramos que *AB* es igual a *CD* querríamos decir que *AB* y *CD* son uno y el mismo segmento. Podríamos, por supuesto, expresar el hecho de que *AB* y *CD* tienen la misma longitud escribiendo:

La longitud de $AB =$ a la longitud de *CD*.

Está claro, del significado dado a '=' que todo es idéntico consigo mismo, esto es, si x es una cosa cualquiera, entonces $x = x$. Este principio recibe algunas veces el nombre de *principio de identidad*.

Además, para cualesquiera x y y, si $x = y$, entonces $y = x$. Este hecho se expresa algunas veces diciendo que la relación de identidad es *simétrica*.

Notamos también que si $x = y$ y $y = z$, entonces $x = z$. Este hecho se expresa diciendo que la relación de identidad es *transitiva*.

Si $x = y$, entonces cualquier cosa que sea verdad respecto de x, es también verdad respecto de y, y cualquier cosa que sea verdad

* La expresión "idéntico a" no debe entenderse aquí en el sentido de perfecta igualdad, sino en el de la misma persona o cosa que. (N. del T.)

respecto de y es también verdad respecto de x. Así, de (4), juntamente con el hecho de que James Madison nació en Port Conway, Virginia, se puede concluir que el cuarto presidente de los Estados Unidos nació en Port Conway, Virginia. Este principio, llamado a menudo el *principio de extensionalidad,* se expresa algunas veces diciendo que los iguales pueden ser sustituidos por iguales.[1]

Si toda propiedad de x es también una propiedad de y, entonces $x = y$, pues x tiene la propiedad de ser idéntica a x y de aquí, que si toda propiedad de x es una propiedad de y, entonces y tiene la propiedad de ser idéntica a x, de manera que $y = x$ y de aquí que $x = y$. Este principio se llama, algunas veces, *ley de Leibniz,* o *principio de la identidad de los indescernibles.*

No hemos, hasta ahora, introducido ningunas reglas especiales que nos permitan obtener *inferencias lógicas* que impliquen identidades. Podríamos, por supuesto, adoptar el punto de vista de que la relación de identidad no tiene una naturaleza especial, en contraste con cualquier otra relación, y que cuando las deduciones implican identidades, se introducirían a medida que se necesiten premisas especiales, como el principio de extensionalidad. Es más conveniente, sin embargo, a causa del ubicuo carácter de las identidades (las usamos en todo dominio sistemático de investigación) introducir una regla especial aplicable a las mismas.

La regla se divide en dos partes. La primera parte simplemente convierte el principio de extensionalidad en un principio de inferencia. Como lo sugiere el ejemplo relativo a James Madison, queremos que la regla sea aplicable no solamente a variables, sino a términos en general. Un ejemplo algebraico pondrá más de relieve esta cuestión. Supóngase que tenemos como línea de derivación:

(1) $$x + y = 2,$$

y sabemos que

(2) $$x = y + 3.$$

[1] Esto no es estrictamente cierto, pues también existen contextos llamados no extensionales, en los cuales no pueden hacerse sustituciones. Considérense, por ejemplo, los tres enunciados siguientes:

(5) La estrella de la mañana es idéntica a la estrella de la tarde.
(6) Juan Pérez sabe que la estrella de la mañana es la misma que Venus.
(7) Juan Pérez sabe que la estrella de la tarde es la misma que Venus.

Aquí (5) es verdadera y (7) resulta de (6) sustituyendo estrella de la mañana por estrella de la tarde; pero puede muy bien suceder que los conocimientos astronómicos de Juan Pérez sean tales que (6) sea verdadera y (7) falsa. Sin embargo, en matemáticas y en sus aplicaciones a las ciencias empíricas, no necesitamos hacer uso de contextos no extensionales, tal como no necesitamos conectivos oracionales no aleticofuncionales. La médula de la cuestión es que rechazando contextos no extensionalistas y conectivos no aleticofuncionales, podemos mantener relativamente sencilla nuestra lógica básica y, no obstante, no renunciar a ningún grado de capacidad de expresión, que es esencial para la vasta superestructura de las matemáticas.

Deseamos estar en capacidad de sustituir la variable x de la Ec. (1) por el término '$y + 3$'. Sin una regla que permita esa sustitución, sería difícil resolver los problemas algebraicos más sencillos.

Para evitar largas restricciones relativas a la presencia de cuantificadores, restringimos la aplicabilidad de la regla a fórmulas *abiertas*, esto es, fórmulas que no tienen cuantificadores. La siguiente inferencia falaz muestra la necesidad de restricción en algún sentido.

{1}	(1) $x = y$	x, y P
{2}	(2) $(\exists x)$-$(x = y)$	y P
{1, 2}	(3) $(\exists x)$-$(x = x)$	1, 2 (regla falaz no restringida para las identidades)

La línea (3) es falsa y es fácil encontrar una interpretación tal que (1) y (2) sean verdaderas: sea la variable libre 'x' 'James Madison' y sea 'y' 'el cuarto presidente de los Estados Unidos'. La fuente de la dificultad en este caso consiste en que cuando 'x' es puesta en (2) en lugar de 'y', con apoyo en (1), 'x' es capturada por el cuantificador existencial ya presente en (2).

La segunda parte de la regla de las identidades simplemente enuncia que el principio de identidad es independiente de cualesquiera premisas, esto es, para cualquier término **t**, la aserción $\mathbf{t} = \mathbf{t}$ es derivable del *conjunto vacío* de premisas. En una deducción indicamos que una línea no depende de ninguna premisa escribiendo una lambda mayúscula, 'Λ', a la izquierda. (Esta letra a menudo se usa para indicar el conjunto vacío, esto es, el conjunto que carece de miembros). Sigue el enunciado formal de la regla.

> REGLA QUE RIGE LAS IDENTIDADES: *Si* **S** *es una fórmula abierta, de* **S** *y* $\mathbf{t}_1 = \mathbf{t}_2$, *o de* **S** *y* $\mathbf{t}_2 = \mathbf{t}_1$ *es posible deducir* **T**, *siempre que* **T** *resulte de* **S** *por reemplazo de una o más incidencia de* $\mathbf{t}_1$, *en* **S**, *por* $\mathbf{t}_2$. *Además, la identidad* $\mathbf{t} = \mathbf{t}$ *es deducible del conjunto vacío de premisas.*

La primera parte de la regla, que es más ampliamente usada en la inferencia lógica que la segunda, se aplica en el siguiente ejemplo:

> EJEMPLO 1. *El hombre que cometió el crimen estaba en el apartamiento. Ahora bien, si alguien estaba en el apartamiento, estaba en la ciudad. Si alguien estaba en México él no estaba en la ciudad. Es un hecho que Barnes estaba en México. En consecuencia, Barnes no es el hombre que cometió el crimen.*

{1}	(1) Ac	P
{2}	(2) $(x)(Ax \rightarrow Tx)$	P
{3}	(3) $(x)(Mx \rightarrow -Tx)$	P
{4}	(4) Mb	P
{3}	(5) $Mb \rightarrow -Tb$	3 *US*
{3, 4}	(6) $-Tb$	4, 5 T
{2}	(7) $Ab \rightarrow Tb$	2 *US*
{2, 3, 4}	(8) $-Ab$	6, 7 T
{9}	(9) $b = c$	P
{2, 3, 4, 9}	(10) $-Ac$	8, 9 I (Regla que rige las identidades)
{1, 2, 3, 4, 9}	(11) $Ac \,\&\, -Ac$	1, 10 T
{1, 2, 3, 4}	(12) $b \neq c$	9, 11 R.A.A.

En la línea (10) se especifica el uso de la regla recién introducida de las identidades. Esta deducción ilustra un método de ataque que es usualmente sano: si deseamos establecer la negación de una identidad, es natural usar una prueba indirecta. Nótese que la aplicación de la regla rectora de las identidades generalmente implica dos líneas previas [(8) y (9) en el ejemplo], a saber, la línea que enuncia una identidad y la que a su vez expone una fórmula en que se hace la sustitución de un término por medio de la identidad. La línea resultante depende, naturalmente, de cualesquiera premisas de que dependen estas líneas. Por ejemplo, la línea (10) depende de las premisas introducidas en las líneas (2), (3), (4) y (9).

En los contextos matemáticos aparecen muy a menudo identidades, juntamente con símbolos de operación. Son típicas fórmulas como:

$$(x)(y)(\exists z)(x + y = z)$$
$$x > 0 \rightarrow (\exists y)(x + y = 0)$$

Los símbolos binarios de operaciones de suma y multiplicación son muy frecuentes, pero el símbolo de operación binario de resta y el símbolo *unario** de operación para encontrar el negativo de un número también aparecen frecuentemente y tienen significados conocidos por todos los lectores de este libro:

$$x - y = 1,$$
$$x + (-x) = 0.$$

Para ilustrar cómo se usan las reglas de inferencia para identidades juntamente con símbolos de operación, podemos deducir algu-

* Se trata de un neologismo que viene desde el inglés (unary); es decir, que el silogismo ha sido creado en inglés y ha resultado necesario pasarlo lexicológicamente intacto, puesto que el morfema inglés -ry corresponde exactamente al castellano -río. (N. del T.)

nas consecuencias de uno de los más sencillos, pero al mismo tiempo de los más importantes conjuntos de axiomas de la matemática moderna, a saber, los axiomas relativos a un *grupo*. Estos axiomas usan tres símbolos no lógicos: el símbolo de operación binaria, '$\circ$', el símbolo de operación unaria '$^{-1}$', y la constante individual 'e'. La interpretación más familiar es la de tomar como dominio de interpretación el conjunto de enteros (positivos, negativos y cero), para interpretar '$\circ$' como '+', '$^{-1}$' como el símbolo de operación para negar un número y 'e' como el nombre de cero. Los tres axiomas para un grupo son:

AXIOMA (1). $(x)(y)(z)(x \circ (y \circ z) = (x \circ y) \circ z)$.
AXIOMA (2). $(x)(x \circ e = x)$.
AXIOMA (3). $(x)(x \circ x^{-1} = e)$.

Así, en la interpretación aritmética que se acaba de proponer, los tres axiomas tienen como interpretación las tres oraciones verdaderas siguientes:

(1') $(x)(y)(z)(x + (y + z) = (x + y) + z)$.
(2') $(x)(x + 0 = x)$.
(3') $(x)(x + (-x) = 0)$.

En la terminología matemática usual, el Axioma (1) dice que la operación $\circ$ es *asociativa;* el Axioma (2) dice que e es un elemento de *identidad* a derecha, con respecto de la operación $\circ$; y el Axioma (3) dice que cada elemento del grupo tiene un *elemento inverso* con respecto de la operación $\circ$. Así, cero es el elemento de identidad para la adición de enteros, y el negativo de un número es su inverso con respecto de la adición.

Pasamos ahora a deducir de los axiomas lo que se conoce como *ley de cancelación a mano derecha*: $(x)(y)(z)(x \circ z = y \circ z \rightarrow x = y)$. Nótese que en la deducción se usan los tres axiomas.

{1}	(1) $(x)(y)(z)(x \circ (y \circ z)$ $= (x \circ y) \circ z)$	P (Ax. (1))
{2}	(2) $(x)(x \circ e = x)$	P (Ax. (2))
{3}	(3) $(x)(x \circ x^{-1} = e)$	P (Ax. (3))
Λ	(4) $(x \circ z) \circ z^{-1}$ $= (x \circ z) \circ z^{-1}$	I
{5}	(5) $x \circ z = y \circ z$	x, y, z P
{5}	(6) $(x \circ z) \circ z^{-1}$ $= (y \circ z) \circ z^{-1}$	x, y, z 4, 5 I
{1}	(7) $x \circ (z \circ z^{-1})$ $= (x \circ z) \circ z^{-1}$	1 *US* $x/x, z/y, z^{-1}/z$

{1}	(8) $y \circ (z \circ z^{-1})$ $= (y \circ z) \circ z^{-1}$	1 *US* $y/x, z/y, z^{-1}/z$
{1, 5}	(9) $x \circ (z \circ z^{-1})$ $= (y \circ z) \circ z^{-1}$	x, y, z 6, 7 I
{1, 5}	(10) $x \circ (z \circ z^{-1})$ $= y \circ (z \circ z^{-1})$	x, y, z 8, 9 I
{3}	(11) $z \circ z^{-1} = e$	3 *US*
{1, 3, 5}	(12) $x \circ e = y \circ e$	x, y 10, 11 I
{2}	(13) $x \circ e = x$	2 *US*
{1, 2, 3, 5}	(14) $x = y \circ e$	x, y 12, 13 I
{2}	(15) $y \circ e = y$	2 *US*
{1, 2, 3, 5}	(16) $x = y$	x, y 14, 15 I
{1, 2, 3}	(17) $x \circ z = y \circ z \rightarrow$ $x = y$	5, 16 C.P.
{1, 2, 3}	(18) $(x)(y)(z)(x \circ z$ $= y \circ z \rightarrow x = y)$	17 *UG*

Las siete aplicaciones en esta deducción de la regla sobre las identidades indica con qué frecuencia se echa mano de ella cuando se manejan símbolos de operación. En el Cap. 7 se desarrollan métodos para reducir la longitud de las deducciones, pero hasta entonces todas deberán ser escritas completas. El ejemplo anterior hace evidente la necesidad de distinguir los pasos de rutina de los que no lo son. Probablemente el paso decisivo en la deducción de la ley de cancelación a mano derecha es el de acertar con la sustitución adecuada en el Axioma (1) [línea (7)], pues esta acuidad de visión lleva consigo la percepción de que $z \circ z^{-1} = e$, en virtud del Axioma (3). El resto de la deducción es sólo cuestión de eliminar el elemento de identidad mediante el uso del Axioma (2).

Volvamos a los axiomas relativos a grupos del §5.2.[2]

Puesto que hemos hecho a la lógica de la identidad parte de nuestra lógica básica de predicados de primer orden (llamada a menudo con esta ampliación: *lógica de predicados de primer orden, con identidad*), necesitamos hacer una observación sobre la extensión a fórmulas que implican identidades, del concepto de interpretación. La extensión es sencilla: el predicado de identidad, como todas las otras constantes lógicas, y a diferencia de todos los otros predicados, se mantiene fijo en su significado original, en todas las interpretaciones. Así, dada la fórmula:

$$(1) \qquad x = 0 \;\&\; y = 0 \rightarrow x + y = 0,$$

[2] Para una clara y detallada, aunque elemental discusión de varios conjuntos de axiomas relativos a grupos, véase *Introduction to Logic*, de Alfred Tarski, Nueva York, 1941.

sería un error interpretar '=' como, digamos, 'es mayor que', en la misma forma que lo sería interpretar '→' como '&'. Esto significa que en fórmulas como la (1) el signo de identidad se trata como un símbolo lógico universal y no como '+', solamente como un símbolo aritmético.[3]

Por otra parte, al dar interpretaciones de predicados no lógicos, es a veces conveniente interpretar un predicado como el predicado lógico de identidad. En realidad, ya se hizo esto en una interpretación del predicado 'I' de indiferencia, dada en el Cap. 4.

EJERCICIOS

1. ¿En cuáles de las siguientes oraciones se usa el verbo 'ser' en el sentido de identidad?

(a) Las mujeres son maravillosas.
(b) Washington era un americano.
(c) Washington fue el primer presidente de los Estados Unidos.
(d) Washington fue el último esposo de Martha Washington.
(e) Jefferson no fue el segundo presidente de los Estados Unidos.
(f) Los hombres no son ángeles.
(g) Simone de Beauvoir no es una gran escritora.
(h) Stephen Dedalus no fue un amante de Molly Bloom.
(i) Stendhal no fue el autor de "Northanger Abbey".

2. Algunos miembros del equipo de natación no han perdido una carrera. Jones pertenece al equipo de natación; además, es el más rápido del equipo. En consecuencia, Jones no ha perdido una carrera. (Mx, Lx, j, f)

3. ¿Es consistente el siguiente conjunto de premisas? Adams es el hombre que firmó el contrato, y el hombre que firmó el contrato está jurídicamente obligado. Sin embargo, Adams está en quiebra y si alguien está en quiebra, entonces no está jurídicamente obligado. (Lx, Bx, a, c)

4. ¿Es consistente el siguiente conjunto de premisas? Horacio es el hombre más valiente del país, pero no es el más fuerte. Pero sólo los valientes son fuertes. (j, b, s, Bx, Sx)

5. Este ejercicio se relaciona con los axiomas relativos a un grupo, pero no depende directamente de ellos. Dado que la operación o es conmutativa y satisface la ley de cancelación a la derecha, pruébese que satisface la ley de cancelación a la izquierda, esto es, de:

(1) $(x)(y)(x \circ y = y \circ x)$,
(2) $(x)(y)(z)(x \circ z = y \circ z \rightarrow x = y)$,

dedúzcase:

$$(x)(y)(z)(z \circ x = z \circ y \rightarrow x = y).$$

[3] La obra de consulta clásica sobre lógica de la identidad es *Grundlagen der Mathematik*, de D. Hilbert y P. Bernays, Vol. 1, Págs. 164 a 209. Puede demostrarse que la lógica de predicados de primer orden con identidad, es completa en el sentido definido en el §4.2.

6. Dados los siguientes tres axiomas relativos a una operación binaria o, pruébese que los axiomas son inconsistentes, o dése una interpretación para demostrar que son consistentes. (Este ejercicio no depende de los axiomas para grupos.)

AXIOMA 1. $(x)(y)(x \circ y = y \circ x)$.
AXIOMA 2. $(x)(y)(x \circ y = y)$.
AXIOMA 3. $(x)(\exists y)(x \neq y)$.

7. Pruébese por el método de interpretación que el primer axioma para grupos es independiente de los otros dos. (Advertencia: no se viole la Regla VII,§4.2).

8. Pruébese por el método de interpretación que el tercer axioma para grupos es independiente de los otros dos. (Advertencia: no se viole la Regla VII,§4.2).

§5.2 Teoremas de lógica. Hasta ahora nos hemos ocupado principalmente de la deducción de conclusiones, a partir de premisas dadas. Deseamos ahora estudiar brevemente fórmulas que son deducibles del conjunto vacío de premisas. Esas fórmulas se llaman *teoremas de lógica,* puesto que son verdaderamente independientes de la verdad o falsedad de cualesquiera premisas fácticas particulares. Haciendo referencia a los conceptos del § 4.2, puede demostrarse que una fórmula es un teorema de lógica si y sólo si es universalmente válido. Puesto que un teorema de lógica no depende de premisas, la estrategia de la prueba es ligeramente diferente a la aplicada para deducir conclusiones de premisas dadas. Si es una implicación la fórmula que estamos intentando establecer como teorema, lo natural es usar el antecedente de la implicación como premisa inicial. En muchos casos puede obtenerse una segunda premisa utilizando como tal negación del consecuente de la implicación y emprendiendo una prueba indirecta. El método puede ilustrarse probando el teorema '$(\exists x)Fx \rightarrow -(x)-Fx$'.

{1}	(1) $(\exists x)Fx$	P
{2}	(2) $(x)-Fx$	P
{1}	(3) $F\alpha$	1 *ES*
{2}	(4) $-F\alpha$	2 *US*
{1, 2}	(5) $F\alpha \,\&\, -F\alpha$	3, 4 T
{1}	(6) $-(x)-Fx$	2, 5 R.A.A.
Λ	(7) $(\exists x)Fx \rightarrow -(x)-Fx$	1, 6 C.P.

Como primera premisa usamos el antecedente '$(\exists x)Fx$' y como segunda premisa '$(x)-Fx$', que es la negación del consecuente. El teorema queda establecido mostrando en la línea (7) que depende del conjunto vacío de premisas, en otras palabras, de ninguna premisa

en absoluto.[4] Así, (7) es verdadera, independientemente de las interpretaciones que se den al predicado 'F'.

Algunas veces resulta necesario adoptar una estrategia más sutil, como en la prueba de que '$-(x)-Fx \rightarrow (\exists x)Fx$' es un teorema.

{1}	(1) Fx	x P
{1}	(2) $(\exists x)Fx$	1 *EG*
Λ	(3) $Fx \rightarrow (\exists x)Fx$	1, 2 C.P.
{4}	(4) $-(\exists x)Fx$	P
{4}	(5) $-Fx$	3, 4 T
{4}	(6) $(x)-Fx$	5 *UG*
Λ	(7) $-(\exists x)Fx \rightarrow (x)-Fx$	4, 6 C.P.
Λ	(8) $-(x)-Fx \rightarrow (\exists x)Fx$	7 T

En este caso no tendría objeto usar el antecedente '$-(x)-Fx$', pues el signo de negación con que comienza esta oración, nos impide hacer algo con ella. En la línea (4) utilizamos como premisa la negación de la conclusión, pero en este caso no para los fines de una prueba indirecta. La principal treta de esta prueba consiste en que aplica *EG* a 'Fx' para obtener (2). Como ya se hizo notar, esas aplicaciones de *EG*, cuando no figura ningún nombre ambiguo, no son frecuentes, pero llegan en ocasiones a ser decisivas.

Los dos teoremas que se acaban de probar están estrechamente relacionados con la Regla Q2 introducida en el § 4.4; el aspecto interesante de que debe tomarse nota es que no se usa Q2 en sus pruebas. Estos dos teoremas son de hecho la base para demostrar que Q2 puede deducirse de las reglas oracionales y de las cuatro reglas introducidas en el Cap. 4.

Puesto que un teorema de lógica ya establecido puede ser útil para probar uno nuevo, es deseable contar con una regla deducida que permita usar directamente el teorema anteriormente establecido, en la deducción que establece el nuevo teorema. Puesto que el uso de resultados previamente establecidos no necesita limitarse a teoremas de lógica, enunciamos en forma general la regla deducida.

Como podía esperarse, se requiere una restricción relativa a los nombres ambiguos, pues el correlato de un nombre ambiguo, esto es, el individuo a que se refiere, cambia de deducción. Considérese, por ejemplo, la deducción:

{1}	(1) $(\exists x)Fx$	P
{1}	(2) $F\alpha$	1 *ES*
Λ	(3) $(\exists x)Fx \rightarrow F\alpha$	1, 2 C.P.

Puesto que la línea (3) es un teorema de lógica, puede ser que

[4] Se explicó en el apartado anterior que en este libro el símbolo 'Λ' designa el conjunto vacío. Nótese su uso en la línea (7).

tengamos la idea de que puede introducirse en cualquir punto, sin crear dificultades. Pero la inferencia falaz de '$(\exists x)Fx$' a '$(x)Fx$' se construye fácilmente usando la línea (3) anterior y una derivación similar y que concierne a '$-Fx$'.

{1}	(1) $(\exists x)-Fx$	P
{1}	(2) $-F\alpha$	1 *ES*
Λ	(3) $(\exists x)-Fx \rightarrow -F\alpha$	1, 2 C.P.
Λ	(4) $F\alpha \rightarrow -(\exists x)-Fx$	3 T
Λ	(5) $(\exists x)Fx \rightarrow F\alpha$	Teorema anterior (falazmente)
Λ	(6) $(\exists x)Fx \rightarrow -(\exists x)-Fx$	4, 5 T
Λ	(7) $(\exists x)Fx \rightarrow (x)Fx$	6 Q1

La dificultad viene de la aparición del nombre ambiguo 'α' en la línea (5), en virtud de la introducción del teorema precedente. Si ambas deducciones se hubieran combinado en una, no se hubiera llegado a la línea (7), pues en la línea (2) de la segunda derivación hubiera tenido que ser introducida 'β' o algún otro nombre ambiguo distinto de 'α' y se hubiera cerrado el paso a la deducción de (6). Se hizo notar, cuando se trató por primera vez de los nombres ambiguos, que éstos eran para uso inmediato en el contexto. Este ejemplo muestra en particular que no pueden ser llevados de deducción a deducción.

REGLA DEDUCIDA PARA INTRODUCIR RESULTADOS PREVIOS. *Si la fórmula* **S** *es deducible de la fórmula* **T** *y de un conjunto de premisas* $\mathscr{P}$ *si* **T** *es deducible de* $\mathscr{P}$*, entonces* **S** *es deducible de* $\mathscr{P}$ *solamente, siempre que* **T** *no contenga nombres ambiguos.*

Si **S** es un teorema de lógica, entonces, por supuesto, $\mathscr{P} = \Lambda$. No daremos una prueba detallada de esta regla deducida, pero resulta evidente tal como está. Por hipótesis, **S** es deducible de **T** y $\mathscr{P}$, y **T** es deducible de $\mathscr{P}$. Para demostrar que **S** es tan sólo deducible de $\mathscr{P}$ simplemente construimos una nueva deducción para **S**, tal que la primera parte de la nueva deducción sea justamente una deducción de **T**, de $\mathscr{P}$.

Para ilustrar el ahorro que esta regla significa, deduzcamos el teorema '$(\exists x)Fx \leftrightarrow -(x)-Fx$'.

Λ	(1) $(\exists x)Fx \rightarrow -(x)-Fx$	Teorema anterior de lógica
Λ	(2) $-(x)-Fx \rightarrow (\exists x)Fx$	Teorema anterior de lógica
Λ	(3) $(\exists x)Fx \leftrightarrow -(x)-Fx$	1, 2 T

Sin la regla, hubiera sido necesario incluir deducciones de (1) y de (2), que habrían aumentado grandemente el número de líneas. La prueba de este teorema ilustra una estrategia típica y normal cuando se prueba un teorema cuyo principal conectivo oracional es 'si y sólo si', subdivimos la prueba en la deducción de dos implicaciones.

Los teoremas de lógica cuyas pruebas se dan como ejercicios ejemplifican todos útiles principios lógicos, algunos de los cuales se discutirán en el apartado siguiente.

El uso de esta regla deducida, para la introducción de resultados previos que no son teoremas de lógica puede ser ilustrada volviendo a los axiomas sobre grupos dados en el apartado anterior. Allí probamos que los axiomas para grupos implican la ley de cancelación a mano derecha, que puede ser enunciada como:

TEOREMA 1 SOBRE TEORÍA DE GRUPOS.

$$(x)(y)(z)(x \circ z = y \circ z \rightarrow x = y).$$

Y podemos usar este teorema para probar un segundo, relativo a la conmutatividad del elemento de identidad.

TEOREMA 2 SOBRE TEORÍA DE GRUPOS. $(x)(x \circ e = e \circ x)$.

DEDUCCION

{1}	(1) $(x)(y)(z)(x \circ (y \circ z) = (x \circ y) \circ z)$	P (Ax. 1)
{2}	(2) $(x)(x \circ e = x)$	P (Ax. 2)
{3}	(3) $(x)(x \circ x^{-1} = e)$	P (Ax. 3)
{1}	(4) $e \circ (x \circ x^{-1}) = (e \circ x) \circ x^{-1}$	1 *US* e/x, x/y, x^{-1}/z
{3}	(5) $x \circ x^{-1} = e$	3 *US*
{1, 3}	(6) $e \circ e = (e \circ x) \circ x^{-1}$	4, 5 I
{2}	(7) $e \circ e = e$	2 *US*
{1, 2, 3}	(8) $e = (e \circ x) \circ x^{-1}$	6, 7 I
{1, 2, 3}	(9) $x \circ x^{-1} = (e \circ x) \circ x^{-1}$	5, 8 I
{1, 2, 3}	(10) $(x)(y)(z)(x \circ z = y \circ z \rightarrow x = y)$	Th. 1
{1, 2, 3}	(11) $x \circ x^{-1} = (e \circ x) \circ x^{-1} \rightarrow x = e \circ x$	10 *US* x/x, $(e \circ x)/y$, x^{-1}/z
{1, 2, 3}	(12) $x = e \circ x$	9, 11 T
{2}	(13) $x \circ e = x$	2 *US*
{1, 2, 3}	(14) $x \circ e = e \circ x$	12, 13 I
{1, 2, 3}	(15) $(x)(x \circ e = e \circ x)$	14 *UG*

Nótese que el Teorema 1 se usa para justificar la introducción de

la línea (10). Es obvio que hubiera sido tedioso e inútil volver a deducir el Teorema 1 en este punto. La regla para introducir resultados previos permite su introducción sin esa repetición. Algunos de los ejercicios tratan de otros teoremas elementales sobre la teoría de los grupos y contienen repetidas aplicaciones de la regla para introducir resultados previos. Al hacer estos ejercicios, debe hacerse un borrador de la línea de ataque, antes de intentar la deducción formal. En un primer intento de poner en marcha una línea de argumentación, es preferible no preocuparse mucho por el rigor lógico, sino experimentar libremente con una variedad de aplicaciones de los axiomas y los teoremas ya probados. En la deducción del Teorema 2 los dos aciertos decisivos consisten en la sustitución que debe hacerse en el Axioma (1), línea (4), y en cómo hacer uso del Teorema 1, línea (11).

EJERCICIOS

1. Pruebe que las siguientes fórmulas son teoremas. Puede usar fórmulas anteriores en la deducción de fórmulas posteriores.

(a) $(x)Fx \rightarrow -(\exists x)-Fx$ (No se use Q1 en la prueba)
(b) $-(\exists x)-Fx \rightarrow (x)Fx$ (No se use Q1 en la prueba)
(c) $(x)(y)Gxy \rightarrow (y)(x)Gxy$
(d) $(\exists x)(\exists y)Gxy \rightarrow (\exists y)(\exists x)Gxy$
(e) $(\exists x)(y)Gxy \rightarrow (y)(\exists x)Gxy$
(f) $(x)(Fx \,\&\, Hx) \rightarrow (x)Fx \,\&\, (x)Hx$
(g) $(x)Fx \,\&\, (x)Hx \rightarrow (x)(Fx \,\&\, Hx)$
(h) $(x)Fx \vee (x)Hx \rightarrow (x)(Fx \vee Hx)$
(i) $(x)(Fx \rightarrow Hx) \rightarrow [(\exists x)Fx \rightarrow (\exists x)Hx]$
(j) $(\exists x)(Fx \,\&\, Hx) \rightarrow (\exists x)Fx \,\&\, (\exists x)Hx$
(k) $Hy \,\&\, (\exists x)Fx \rightarrow (\exists x)(Hy \,\&\, Fx)$
(l) $(\exists x)(Fx \vee Hx) \leftrightarrow (\exists x)Fx \vee (\exists x)Hx$
(m) $(x)(Hy \vee Fx) \leftrightarrow Hy \vee (x)Fx$
(n) $(\exists x)(Hy \vee Fx) \leftrightarrow Hy \vee (\exists x)Fx$
(o) $(x)(Hy \,\&\, Fx) \leftrightarrow Hy \,\&\, (x)Fx$
(p) $(\exists x)(Hy \,\&\, Fx) \leftrightarrow Hy \,\&\, (\exists x)Fx$
(q) $(x)(Hy \rightarrow Fx) \leftrightarrow [Hy \rightarrow (x)Fx]$
(r) $(\exists x)(Hy \rightarrow Fx) \leftrightarrow [Hy \rightarrow (\exists x)Fx]$
(s) $(x)(Fx \rightarrow Hy) \leftrightarrow [(\exists x)Fx \rightarrow Hy]$
(t) $(\exists x)(Fx \rightarrow Hy) \leftrightarrow [(x)Fx \rightarrow Hy]$

2. Para cada una de las siguientes fórmulas pruébese por medio de una deducción o de una interpretación, según el caso, si la fórmula es (1) un teorema de lógica, (2) la negación de un teorema de lógica, o (3), ni (1) ni (2).

(a) $(x)(y)(z)((Exy \,\&\, Fyz) \rightarrow Fxz)$
(b) $(x)(y)(Fx \vee -Fy)$
(c) $(x)(\exists y)(Fx \,\&\, Gy) \rightarrow (\exists y)(x)(Fx \,\&\, Gy)$
(d) $(x)(Fx \leftrightarrow -Fx) \rightarrow (\exists y)Fy$

(e) $(x)(\exists y)(Fx \,\&\, -Fx \,\&\, -Fy)$
(f) $(\exists x)(Fx \vee -Fx)$
(g) $(x)Fx \vee (\exists y)-Fy$

3. Pruébese que las cinco fórmulas siguientes relativas a la relación de identidad son teoremas de lógica.

(a) $(x)(y)(x = y \rightarrow y = x)$
(b) $(x)(y)(z)[(x = y \,\&\, y = z) \rightarrow x = z]$
(c) $(x)(y)(z)[(x = z \,\&\, y = z) \rightarrow x = y]$
(d) $(x)(\exists y)(x = y)$
(e) $(x)[Fx \leftrightarrow (\exists y)(x = y \,\&\, Fy)]$

4. Este ejercicio trata de la teoría elemental de grupos. Usando los axiomas sobre grupos y los dos teoremas ya establecidos, formúlense deducciones formales de los teoremas que siguen.

TEOREMA 3. (CARÁCTER ÚNICO DE LA IDENTIDAD).

$$(y)[(x)(x \circ y = x) \rightarrow y = e]$$

TEOREMA 4. (CONMUTATIVIDAD DEL INVERSO).

$$(x)(x \circ x^{-1} = x^{-1} \circ x)$$

TEOREMA 5. (LEY DE CANCELACIÓN A MANO IZQUIERDA).

$$(x)(y)(z)(z \circ x = z \circ y \rightarrow x = y)$$

TEOREMA 6. (CARÁCTER ÚNICO DEL INVERSO).

$$(x)(y)(x \circ y = e \rightarrow y = x^{-1})$$

TEOREMA 7. $(x)((x^{-1})^{-1} = x)$
TEOREMA 8. $(x)(y)(\exists z)(x = y \circ z)$
TEOREMA 9. $(x)(z)(\exists y)(x = y \circ z)$

5. Los dos últimos teoremas del ejercicio anterior pueden combinarse con el teorema de la asociatividad para dar un nuevo grupo de axiomas sobre la teoría de grupos, basados solamente en el único símbolo no lógico '$\circ$'; esto es, en esta formulación los axiomas son:

AXIOMA 1. $(x)(y)(z)(x \circ (y \circ z) = (x \circ y) \circ z)$
AXIOMA 2. $(x)(y)(\exists z)(x = y \circ z)$
AXIOMA 3. $(x)(z)(\exists y)(x = y \circ z)$

Para indicar cómo va el desarrollo, demuéstrense los siguientes cuatro teoremas sobre la base de estos tres nuevos axiomas:

TEOREMA 1. (EXISTENCIA DEL ELEMENTO DE IDENTIDAD A MANO DERECHA).

$$(\exists y)(x)(x \circ y = x)$$

TEOREMA 2. (EXISTENCIA DEL ELEMENTO DE IDENTIDAD A MANO IZQUIERDA).

$$(\exists y)(x)(y \circ x = x)$$

TEOREMA 3. (CARÁCTER ÚNICO DEL ELEMENTO DE IDENTIDAD A MANO DERECHA).

$$(y)(z)[(x)(x \circ y = x \,\&\, x \circ z = x) \rightarrow y = z]$$

TEOREMA 4. (CARÁCTER ÚNICO DEL ELEMENTO DE IDENTIDAD A MANO IZQUIERDA).

$$(y)(z)[(x)(y \circ x = x \,\&\, z \circ x = x) \rightarrow y = z]$$

§5.3 Reglas deducidas de inferencia. En el §2.4 introdujimos una regla derivada para pruebas indirectas; se hizo la observación de que podían deducirse las tres reglas del §4.4 y en el apartado anterior se enunció una regla relativa al uso de teoremas. Queremos ahora examinar en forma más general el carácter de esas reglas. Podemos abordar este problema preguntando en qué difieren las reglas deducidas de inferencia de reglas originales como la regla de especificación universal o la de la prueba condicional. Hay cuando menos tres puntos esenciales de diferencia.

(i) Las reglas, derivadas de inferencia no son indispensables; desde un punto de vista teórico son un simple expediente cómodo. Cualquier inferencia que usa una regla deducida puede ser siempre desplazada por otra que no la use. Por ejemplo, en el caso de una regla aplicable a pruebas indirectas resulta claro de la discusión contenida en el §2.4 que eliminar su uso es una cuestión sencilla y trivial.

(ii) Las reglas derivadas de inferencia deben ser *demostradas* con apoyo en las reglas originales. En el presente contexto una prueba consiste primordialmente en dar un conjunto explícito de instrucciones para eliminar el uso de la regla en cualquier ocasión particular. El punto de vista es que una regla deducida puede aplicarse en la inferencia, si es posible formular para un escéptico Santo Tomás un conjunto de instrucciones explícitas relativas a una derivación formal que no use la regla deducida, y eso no obstante sea apta para llegar a la misma conclusión. La prueba de la Regla R.A.A. dio ese conjunto de instrucciones. En esa prueba la idea fundamental es que podemos suprimir la referencia a la regla aplicando la Ley de Reducción al Absurdo.

(iii) Las reglas originales se usan para caracterizar en forma *efectiva* o *finitística** el concepto intuitivo de inferencia lógica. Como

* En una nota anterior hemos hecho referencia a un neologismo que procedía del inglés, esto es, había sido creado en ese idioma; se trata aquí de un caso igual. Por lo demás, la necesidad del vocablo "finitístico" resulta evidente si se tienen en cuenta los peligros de confusión que contrae el uso de otras palabras como, por ejemplo, "teleológico". (N. del T.)

vimos en el §4.2, el concepto de argumento válido o de oración universalmente válida se define en función del concepto de interpretación verdadera de una fórmula e independientemente de *cualesquiera* reglas de inferencia. Sin embargo, por lo común no es práctico demostrar que un argumento es válido por el procedimiento de mostrar que toda interpretación que hace verdaderas a las premisas hace también verdadera a la conclusión. Puesto que hay un número infinito de interpretaciones de un argumento, es literalmente imposible examinar todas las interpretaciones y determinar su verdad o falsedad para las premisas y conclusiones. De este modo surge naturalmente la idea de desarrollar un número finito de reglas de inferencia que pueda usarse para establecer la validez de los argumentos. Las dos propiedades importantes que deben tener las reglas originales son las de legitimidad o solidez y la de totalidad (ya discutidas en los Caps. 2 y 4). Las reglas deducidas, como son teóricamente redundantes, no entran directamente en la prueba de la legitimidad o la totalidad de las reglas básicas de inferencia, aunque la prueba de una nueva regla derivada debería mostrar que la nueva regla no puede convertir un sistema legítimo de inferencia en uno ilegítimo.

Hasta aquí por lo que hace a reglas de inferencia originales y deducidas. Para todo teorema de lógica hay una correspondiente regla derivada. Puesto que hay un número ilimitado de teoremas, hay un número ilimitado de reglas derivadas de inferencia. Sin embargo, es relativamente pequeño el número de esas reglas que se usan repetidamente en la inferencia. Limitaremos el resto de este apartado al estudio de algunas reglas más que gobiernan el uso de cuantificadores y al de una extensión de la regla relativa a equivalencias tautológicas (que se enunció en el §4.4) a otra regla para intercambiar *lógicamente* fórmulas equivalentes.

Si deseáramos establecer Q1 como regla derivada, sería razonable comenzar por probar que:

(A) De $(\mathbf{v})\mathbf{S}$ podemos deducir $-(\exists \mathbf{v})-\mathbf{S}$.

Demostración: Introduciendo $(\mathbf{v})\mathbf{S}(\mathbf{v})$ y $(\exists \mathbf{v})-\mathbf{S}(\mathbf{v})$ como premisas, deducimos una contradicción (i) usando *ES* para obtener $-\mathbf{S}(\mathbf{v})$ de $(\exists \mathbf{v})-\mathbf{S}(\mathbf{v})$, donde $\mathbf{v}$ es un nombre ambiguo apropiado, y (ii) usando *US* para obtener $\mathbf{S}(\mathbf{v})$ de $(\mathbf{v})\mathbf{S}(\mathbf{v})$. La contradicción $\mathbf{S}(\mathbf{v})$ & $-\mathbf{S}(\mathbf{v})$ establece la regla por prueba indirecta.

Podemos usar la Regla (A) para dar una derivación en dos líneas de '$-(\exists x)-(x+0=x)$', a partir de '$(x)(x+0=x)$'.

{1}	(1) $(x)(x+0=x)$	P
{1}	(2) $-(\exists x)-(x+0=x)$	1, (A)

Si se pusiera en duda la validez de esta deducción, podríamos usar la demostración de la Regla (A) para eliminar la referencia a (A) en la línea (2); tendríamos entonces, en función de las reglas originales, la derivación más larga:

{1}	(1) $(x)(x+0=x)$	P
{2}	(2) $(\exists x)-(x+0=x)$	P
{2}	(3) $-(\alpha+0=\alpha)$	2 *ES*
{1}	(4) $\alpha+0=\alpha$	1 *US*
{1, 2}	(5) $(\alpha+0=\alpha)\ \&\ -(\alpha+0=\alpha)$	3, 4 T
{1}	(6) $-(\exists x)-(x+0=x)$	2, 5 R.A.A.

La función de las reglas derivadas es primordialmente la de eliminar repeticiones innecesarias de patrones recurrentes de inferencia. Una vez que está claro el método general de ataque, no tiene caso duplicar el argumento correspondiente a las líneas (2, a (5), cuando deseamos pasar de un enunciado universal a la negación de un enunciado existencial.

En la regla que sigue se resumen las propiedades más importantes de los cuantificadores; la primera mitad de la primera parte de esta regla es lo que se acaba de discutir con el nombre de Regla (A).

REGLA DEDUCIDA QUE RIGE LOS CUANTIFICADORES: **Q**

(1) *Si* **S** *es una fórmula, entonces*:

(a) *de* (**v**)**S** *se puede deducir* –(∃ **v**)–**S**, *y recíprocamente;*
(b) *de* (**v**)–**S** *se puede deducir* –(∃ **v**)**S**, *y recíprocamente;*
(c) *de* (∃ **v**)**S** *se puede deducir* –(**v**)–**S**, *y recíprocamente;*
(d) *de* (∃ **v**)**S** *se puede deducir* –(**v**)**S**, *y recíprocamente;*
(e) *de* (**v**)(**w**)**S** *se puede deducir* (**w**)(**v**)**S**;
(f) *de* (∃ **v**)(∃ **w**)**S** *se puede deducir* (∃ **w**)(∃ **v**)**S**;
(g) *de* (∃ **v**)(**w**)**S** *se puede deducir* (**w**)(∃ **v**)**S**;

(2) *Si* **S** *y* **T** *son fórmulas cualesquiera, entonces*:

(a) *de* (v)(S & T), *se puede deducir* (v)S & (v)T T *y recíprocamente;*
(b) *de* (∃ **v**)(**S** ∨ **T**), *se puede deducir* (∃ **v**)**S** ∨ (∃ **v**)**T**, *y recíprocamente;*
(c) *de* (**v**)(**S** → **T**) *se puede deducir* (∃ **v**)**S** → (∃ **v**)**T**;
(d) *de* (∃ **v**)(**S** & **T**) *se puede deducir* (∃ **v**)**S** & (∃ **v**)**T**;
(e) *de* (**v**)**S** ∨ (**v**)**T** *se puede deducir* (**v**)(**S** ∨ **T**).

(3) *Si* **S** *y* **T** *son fórmulas cualesquiera y* **v** *no está libre en* **S**, *entonces*:

(a) *de* (**v**)(**S** ∨ **T**), *se puede deducir* **S** ∨ (**v**)**T**, *y recíprocamente;*

(b) *de* (∃**v**)(**S** ∨ **T**) *se puede deducir* **S** ∨ (∃**v**)**T**, y recíprocamente;

(c) *de* (**v**)(**S** & **T**) *se puede deducir* **S** & (**v**)**T**, *y recíprocamente;*

(d) *de* (∃**v**)(**S** & **T**) *se puede deducir* **S** & (∃**v**)**T**, *y recíprocamente;*

(e) *de* (**v**)(**S** → **T**) *se puede deducir* **S** → (**v**)**T**, *y recíprocamente;*

(f) *de* (∃**v**)(**S** → **T**) *se puede deducir* **S** → (∃**v**)**T**, *y recíprocamente;*

(g) *de* (**v**)(**T** → **S**) *se puede deducir* (∃**v**)**T** → **S**, *y recíprocamente.*

(h) *de* (∃**v**)(**T** → **S**) *se puede deducir* (**v**)**T** → **S**, *y recíprocamente.*

En líneas anteriores se dio la demostración de la primera mitad de la parte primera. El resto de la demostración de las diversas partes de esta regla se deja al estudiante. La demostración de casi todas las partes es semejante a la deducción de uno o dos teoremas del Ejercicio 1 del apartado anterior. Por ejemplo, (1a) de la regla corresponde a (a) y (b) del ejercicio; (3a) de la regla corresponde a (m) del ejercicio. La demostración de (1b) y (1d) es trivial una vez que han sido establecidos (1a) y (1c), respectivamente.

Difícilmente se retendrán de una sola vez todas las partes de las reglas anteriores, pero resultará útil tener en algún lugar a la mano, para consulta, una regla tan completa como ésta. Puesto que los estudiantes tienen ocasionalmente dificultades para interpretar correctamente la restricción de la tercera parte de la regla, de que **v** no está libre en **S**, puede ser útil estudiar uno o dos ejemplos. Dada la fórmula '$(x)(Mx \rightarrow Ax)$', ¿puede usarse (3e) para deducir '$Mx \rightarrow (x)Ax$'? No; puesto que 'x' está libre en 'Mx', la restricción no está satisfecha. En este ejemplo, **S** es 'Mx' y **T** es 'Ax'. El hecho de que 'x' esté ligada en toda la fórmula '$(x)(Mx \rightarrow Ax)$' no tiene consecuencias; la cuestión es siempre: ¿Está 'x' libre en **S** considerada en sí misma? Como un ejemplo más, supóngase que se nos da '$(\exists y)[(x)(Mx \rightarrow Ax) \vee Ay]$'. ¿Podemos usar (3b) para deducir '$(x)(Mx \rightarrow Ax) \vee (\exists y)Ay$'? Sí; pues aquí **S** es '$(x)(Mx \rightarrow Ax)$', **v** es la variable 'y' y claramente 'y' está libre en '$(x)(Mx \rightarrow Ax)$'. En este ejemplo **T** es, por supuesto, 'Ay'.

Los usos de la regla deducida que rige los cuantificadores se amplían, extendiendo la regla que permite el reemplazo de una fórmula por una tautológicamente equivalente, hasta permitir el reemplazo por una fórmula *lógicamente* equivalente. No hemos todavía caracterizado con precisión el concepto de equivalencia lógica, pero del §4.2 debe resultar obvia su adecuada caracterización. Dos fórmulas son *lógicamente equivalentes* si y sólo si cada una implica la otra. Esto es lo mismo que decir que la bicondicional formada por ellas es universalmente válida. De la legitimidad y totalidad de las reglas básicas de inferencia se sigue que dos fórmulas son lógicamente equivalentes si y sólo si cada una es deducible de la otra por medio de las reglas de inferencia. O, dicho de otro modo, si la bicondicional formada por las mismas es un teorema de lógica.

Supóngase, por ejemplo, que deseamos demostrar que la fórmula:

(1) $$(x)(\exists y)(xQy \rightarrow (z)(\exists u)(yQz \rightarrow uQx))$$

implica a y es implicada por:

(2) $$(x)(\exists y)(z)(xQy \,\&\, yQz \rightarrow (\exists u)(uQx)).$$

Observamos que se sigue la regla Q3*e* anterior que la fórmula

(3) $$xQy \rightarrow (z)(\exists u)(yQz \rightarrow uQx)$$

es lógicamente equivalente a:

(4) $$(z)(xQy \rightarrow (\exists u)(yQz \rightarrow uQx)).$$

Además, por Q3*f* vemos que la fórmula:

(5) $$(\exists u)(yQz \rightarrow uQx)$$

es lógicamente equivalente a:

(6) $$yQz \rightarrow (\exists u)(uQx).$$

Reemplazando entonces (5) por (6) en (4), obtenemos como lógicamente equivalente a (4) y, en consecuencia, como lógicamente equivalente a 3:

(7) $$(z)(xQy \rightarrow (yQz \rightarrow (\exists u)uQx)),$$

pero la fórmula:

$$xQy \rightarrow (yQz \rightarrow (\exists u)uQx)$$

es tautológicamente equivalente (por importación de 'yQx') a:

(8) $$xQy \,\&\, yQz \rightarrow (\exists u)uQx,$$

de donde (7), y en consecuencia (3), es lógicamente equivalente a:

(9) $(z)(xQy \ \& \ yQz \rightarrow (\exists u)uQx)$.

Reemplazando (3) por (9) en (1) obtenemos (2) inmediatamente, cómo deseábamos. Nótese que Q3*e* y Q3*f* no podrían ser aplicados directamente a (1), puesto que los cuantificadores a que las aplicamos no estaban al frente de la fórmula. Se necesitó la regla que permite el mutuo reemplazo de fórmulas lógicamente equivalentes. Naturalmente, podría haberse demostrado la equivalencia de (1) y (2) por medio de una deducción que no hubiera hecho uso de reglas derivadas, pero esto implica la tediosa operación de suprimir y volver a introducir todos los cuantificadores, que es precisamente lo que se trató de que ayudaran a evitar las reglas, al ser éstas diseñadas.

La regla para el reemplazo mutuo de fórmulas lógicamente equivalentes debe, como la regla para introducir resultados previos, restringirse con respecto de los nombres ambiguos. La construcción de una inferencia falaz que justifique la restricción se deja como ejercicio.

REGLA PARA FÓRMULAS LÓGICAMENTE EQUIVALENTES: L.E. *Sean* **P** *y* **Q** *fórmulas lógicamente equivalentes que no contienen nombres ambiguos. Si* **P** *ocurre como parte de la fórmula* **R** *y si la fórmula* **S** *resulta de* **R** *reemplazando por* **Q** *cuando menos una incidencia de* **P** *en* **R**, *entonces* **S** *es deducible de* **R** *y recíprocamente; esto es,* **R** *y* **S** *son lógicamente equivalentes.*

La legitimidad intuitiva de la regla debe resultar evidente. Puesto que una demostración rigurosa de esta regla implica el uso del principio de inducción matemática a todo lo largo de **R**, se omiten aquí los detalles, pero puede enunciarse el concepto intuitivo. Si **R** es simplemente **P**, la prueba es trivial, pues la equivalencia lógica de **R** y **S** es la misma cosa que la equivalencia lógica de **P** y **Q**. Suponemos ahora que el teorema es válido para fórmulas de longitud no mayor que n y demostramos que también lo es para cualquier fórmula **R** de longitud $n + 1$. Este último argumento se descompone en varios casos. **R** debe ser de la forma $-\mathbf{R}_1$, $\mathbf{R}_1 \ \& \ \mathbf{R}_2$, $\mathbf{R}_1 \vee \mathbf{R}_2$, $\mathbf{R}_1 \rightarrow \mathbf{R}_2$, $\mathbf{R}_1 \leftrightarrow \mathbf{R}_2$, $(\mathbf{v})\mathbf{R}_1$, o $(\exists \mathbf{v})\mathbf{R}_1$, donde $\mathbf{R}_1$ y $\mathbf{R}_2$ son ambos de longitud menor que $n + 1$. El problema es, pues, el de demostrar que en cada uno de los casos es válido para **R** el teorema, en la hipótesis inductiva de que es válido para $\mathbf{R}_1$ y $\mathbf{R}_2$; la prueba es directa y sencilla para cada caso. Supóngase, por ejemplo, que **R** es $\mathbf{R}_1 \ \& \ \mathbf{R}_2$. Sea $\mathbf{S}_1$ el resultado de reemplazar cualquier número de incidencias de **P** en $\mathbf{R}_1$ por **Q**, y supóngase que $\mathbf{S}_2$ resulta de $\mathbf{R}_2$ en forma similar. Enton-

ces, por la hipótesis inductiva, tenemos como lógicamente equivalentes y, en consecuencia, como teoremas de lógica,

(1) $$R_1 \leftrightarrow S_1$$

(2) $$R_2 \leftrightarrow S_2,$$

y (1) y (2) implican tautológicamente

$$R_1 \ \& \ R_2 \leftrightarrow S_1 \ \& \ S_2;$$

esto es, **R** y **S** son lógicamente equivalentes como se deseaba, puesto que **R** es $\mathbf{R}_1$ & $\mathbf{R}_2$, y **S** es $\mathbf{S}_1$ & $\mathbf{S}_2$.

Con respecto a las observaciones hechas en el §4.4 sobre el carácter derivado de las reglas Q1, Q2 y T.E., está ahora claro que las tres se siguen inmediatamente de la regla para los cuantificadores enunciada en este apartado y la regla del reemplazo de fórmulas lógicamente equivalentes.

Con esto hemos completado nuestro desarrollo de las reglas de inferencia lógica. Es característica de la exposición elemental contenida en los Caps. 1 a 5 que no hayamos demostrado nada complicado o refinado *sobre* lógica. La finalidad de esos capítulos ha sido la de enseñar cómo usar exactas y explícitas reglas de inferencia. Los capítulos siguientes (con excepción del que sigue inmediatamente) tratan de la aplicación de los conceptos y métodos de la lógica hasta aquí presentados, al contexto al orden de ideas más amplios de las matemáticas. En otras palabras, nos dedicaremos a aplicar nuestra herramienta lógica a la comprensión de las matemáticas y del carácter de la ciencia teórica en general.

EJERCICIOS

1. En cada uno de los siguientes casos construya una deducción falaz para demostrar que la adopción de la regla particular en cuestión conduciría de premisas verdaderas a una falsa conclusión.

   (a) De $(v)(\exists w)S$ podemos deducir $(\exists w)(v)S$.
   (b) De $(v)(S \rightarrow T)$ podemos deducir $(\exists v)S \rightarrow (v)T$.
   (c) De $(\exists v)S$ & $(\exists v)T$ podemos deducir $(\exists v)(S \ \& \ T)$.
   (d) De S & $(\exists v)T$ podemos deducir $(\exists v)(S \ \& \ T)$. (SUGESTIÓN: Considérese el caso en que v está libre en S).
   (e) De $(v)(S \vee T)$ podemos deducir $(v)S \vee (v)T$.
   (f) De $(v)(S \rightarrow T)$ podemos deducir $(\exists v)S \rightarrow T$.
   (g) De $(v)S \rightarrow (v)T$ podemos deducir $(v)(S \rightarrow T)$.

2. Demuestre los siguientes teoremas de lógica.
   (a) $(x)(Fx \vee -Fx)$
   (b) $Gy \rightarrow (x)(Gy \vee Fx)$
   (c) $(x)(Fx \leftrightarrow Gx) \rightarrow ((x)Fx \leftrightarrow (x)Gx)$
   (d) $-(x)(y)(\exists z)Fxyz \leftrightarrow (\exists x)(\exists y)(z)-Fxyz$

3. Demuestre las diversas partes de las reglas que rigen los cuantificadores.

4. ¿En cuál de lo que sigue se hace una aplicación correcta de la parte indicada de la regla que rige los cuantificadores. Nótese que comúnmente el problema crucial es la decisión si las restricciones en variables libres se han satisfecho.

(a) Según (3a), de '$Fy \vee (x)Fx$' podemos deducir '$(x)(Fy \vee Fx)$'.
(b) Según (3b), de '$Hxy \vee (\exists y)Fy$' podemos deducir '$(\exists y)(Hxy \vee Fy)$'.
(c) Según (3c), de '$(x)(\exists y)(Fy \,\&\, Hxy)$', podemos deducir '$(\exists y)(Fy \,\&\, (x)Hxy)$'.
(d) Según (3d), de '$(\exists x)(x)Fx \,\&\, Hxy)$' podemos deducir '$(x)Fx \,\&\, (\exists x)Hxy$'.
(e) Según (3e), de '$(x)(y \neq 0 \rightarrow (x \neq 0 \rightarrow x \cdot y \neq 0)$', podemos deducir '$y \neq 0 \rightarrow (x)(x \neq 0 \rightarrow x \cdot y \neq 0)$'.
(f) Según \(3f), de '$(\exists y)(x > 1 \,\&\, y > 1 \rightarrow x \cdot y > 1)$', podemos deducir '$(x > 1 \,\&\, y > 1) \rightarrow (\exists y)(x \cdot y > 1)$'.[1]
(g) Según (3g), de '$(z)x + z = y + z \rightarrow x = y)$', podemos deducir '$(\exists z)(x + z = y + z) \rightarrow x = y$'.
(h) Según (3h), de '$(\exists y)(y > 0 \rightarrow x > 0)$' podemos deducir '$(y)(y > 0) \rightarrow x > 0$'.

5. En cada una de las siguientes inferencias se hace uso de la regla para el reemplazo lógico de fórmulas equivalentes (L.E.). En cada caso enúnciese la equivalencia lógica que se use en particular. La fórmula de la derecha se deduce de la de la izquierda y recíprocamente.

(a) $(x)(y)(\exists z)(x > y \rightarrow x > z \,\&\, z > y)$
$(x)(y)(x > y \rightarrow (\exists z)(x > z \,\&\, z > y))$
(b) $(\exists x)(y)(y \geq 0 \rightarrow y > x)$ $(\exists x)(y)(y > x \vee -(y \geq 0))$
(c) $(\exists x)(y)-(y > x)$ $(\exists x)-(\exists y)(y > x)$
(d) $(\exists x)-(y)(x > 0 \rightarrow y - 1 < 2)$ $(\exists x)(\exists y)(x > 0 \,\&\, -(y - 1 < 2))$

6. Demuéstrese la siguiente regla deducida. (Sustitución de Variables Libres): *De* **S(v)** *se puede deducir* **S(*t*)**, *siempre que* (*i*) **v** *no esté imprecisa,* (*ii*) **v** *no sea un subíndice y* (*iii*) *ninguna incidencia libre de* **v** *quede dentro del alcance de un cuantificador que use una variable de t.*

7. Constrúyase una inferencia falaz que justifique la restricción relativa a nombres ambiguos en la regla para reemplazar lógicamente fórmulas equivalentes.

CAPITULO 6

APOSTILLA SOBRE EL USO Y LA MENCION[1]

§6.1. Nombres y cosas nombradas. En el uso común del lenguaje no existe la posibilidad de confundir una cosa con su nombre. Usamos nombres para hablar de las cosas y, al parecer, sólo un idiota podría confundir a Guillermo Shakespeare con su nombre. No obstante, cuando deseamos mencionar nombres o expresiones en general y no solamente usarlos, no necesitamos ser idiotas para sufrir confusiones. Esto es, se suscitan ciertos problemas especiales cuando las cosas nombradas son expresiones lingüísticas. El método tipo para designar expresiones consiste en usar comillas dobles o sencillas; en capítulos anteriores hemos usado comillas sencillas.

Considérense las siguientes oraciones:

(1) California es un estado.
(2) California tiene diez letras.
(3) 'California' es un estado.
(4) 'California' tiene diez letras.
(5) 'California' es un nombre de California.
(6) "California" es un nombre de un nombre de California.

Sobre la base de las observaciones del primer párrafo debe resultar claro que las oraciones (1), (4), (5) y (6) son verdaderas, mientras que (2) y (3) son falsas. En las oraciones (3), (4) y (5) la palabra "California" está siendo mencionada más bien que usada. En tales casos resulta más claro enunciar la expresión 'la palabra' antes de mencionar la palabra de que se trate. Corregida de este modo, la oración (3) se leería:

(3') La palabra 'California' es un estado.

Una prueba más sutil de nuestra comprensión de estas cuestiones

[1] Este capítulo puede omitirse sin pérdida de continuidad.

puede obtenerse bautizando la palabra 'California' con un nombre propio de persona. Sea

(7) Jeremías = 'California'.

Sobre la base de (7) veamos cuál de los siguientes enunciados es verdadero:

(8) Jeremías es un nombre.
(9) Jeremías tiene diez letras.
(10) Jeremías tiene ocho letras.
(11) 'Jeremías' tiene ocho letras.
(12) 'Jeremías' es el nombre de un estado.

Debe resultar evidente que (8), (9) y (11) son verdaderos, mientras que (10) y (12) son falsos. Si en (12) se hubiera usado la palabra Jeremías y no mencionado, se habría obtenido una oración verdadera,

(12′) Jeremías es el nombre de un estado.

La oración podría también haberse vuelto verdadera insertando la expresión adicional 'nombre de':

(12″) 'Jeremías' es el nombre del nombre de un estado.

El punto de vista aceptado por la mayor parte de los filósofos es que no toda expresión lingüística designa algo. Un caso extremo es el del paréntesis a la izquierda. Casi todos convendrían en que esta expresión no es el nombre de un objeto. Se suscitan problemas aún más intrigantes cuando se trata de predicados. Por ejemplo, ¿qué clase de entidad designa, si es que designa alguna, el predicado 'es un presidente de los Estados Unidos, del siglo XIX'? No entraremos aquí en esta controversia, ni son necesarios estos supuestos controvertidos para nuestros subsecuentes desarrollos. La distinción que en este libro resulta de mucho interés es la que existe entre las expresiones y sus nombres. El paréntesis a la izquierda tiene su nombre, aun cuando no designa nada.

EJERCICIOS

1. ¿Cuál de las siguientes oraciones es verdadera?

(a) Kant fue un filósofo alemán.
(b) 'Kant" tiene cuatro letras.
(c) Kant es un nombre de Kant.
(d) Kant es un nombre de 'Kant'.
(e) Newton es un apellido más largo que Kant.
(f) Newton es un nombre más largo que Kant.

2. Si, dando un nombre al nombre, hacemos

María = 'Maruca',

¿cuáles de las siguientes oraciones son verdaderas?

(a) María es un nombre más largo que 'María'.
(b) 'María' tiene siete letras.
(c) María es un nombre común entre las muchachas.
(d) 'María' es un nombre común entre las muchachas.
(e) No hay ninguna persona llamada María.
(f) María es un nombre de Maruca.
(g) 'María' es un nombre de Maruca.
(h) María es un nombre de 'Maruca'.
(i) 'María' es un nombre de 'Maruca'.

§6.2. Problemas de las variables oracionales. En el Cap. 1 se nos presentó el problema de usar variables que podían ser sustituidas por oraciones o nombres de oraciones. Por varias razones se escogió la segunda alternativa; las letras 'P', 'Q', 'R' etc., del Cap. 1 se sustituyen con nombres de oraciones. Cuando se tiene el propósito de reemplazar las variables por oraciones y no por nombres de oraciones, se acostumbra usar las letras minúsculas '*p*', '*q*', '*r*', etc. Sobre la base de las convenciones establecidas, es adecuado llamar a las letras 'P', 'Q' y 'R' *variables oracionales* y las letras '*p*', '*q*' y '*r*' *variables proposicionales*. Ahora bien, la palabra 'proposición' tiene en la filosofía reciente un historial deslucido. Entendemos aquí que una proposición es una entidad nombrada por una oración. Si no se cree o no se quiere creer que las oraciones nombren algo, no se usen variables proposicionales.

La observación que se acaba de hacer sobre las variables proposicionales resultaría sospechosa en algunos reductos filosóficos, pero su base parece legítima al autor y puede explicarse recurriendo a las discusiones anteriores sobre variables y términos en general, contenidas en los Caps. 3 y 4. El punto de vista tipo de la lógica formal es:

(I) *Cuando una variable es reemplazada por un término que no contiene variables, entonces el término reemplazado debe nombrar una entidad.*

Esta tesis, esto es, la tesis (I), es tan fundamental para el desarrollo de una lógica legítima y de fácil marcha, que ha de abreviarse sólo por razones muy profundas. (En el Cap. 8 puede encontrarse una aplicación específica de esta tesis en la discusión de lo que debe hacerse sobre el problema aritmético de dividir por cero). Dada la oración verdadera de aritmética:

$$(x)(y)(x + y = y + x),$$

si inferimos por especificación universal,

$$3 + 5 = 5 + 3,$$

quedamos entonces comprometidos por (I) a adoptar el punto de vista de que '3 + 5' y '5 + 3' son términos que designan cierto número. Las letras '*x*' y '*y*' se llaman *variables numéricas* porque pueden ser reemplazadas por términos que designan números. De modo semejante, de la verdad de la lógica de que para cualesquiera proposiciones p y q

$$p \rightarrow (p \vee q),$$

podemos inferir

$$1 = 1 \rightarrow (1 = 1 \vee 2 = 3).$$

Puesto que la variable '*p*' es reemplazada por la oración '1 = 1', de acuerdo con (I), esta oración debe ser el nombre de alguna entidad. De allí la necesidad de proposiciones cuando se usan variables que son reemplazadas con oraciones y no con nombres de oraciones.[1]

Haciendo a un lado el escepticismo sobre la existencia de entidades tales como las proposiciones, tienen ciertas ventajas técnicas las variables que son reemplazadas por nombres de oraciones y no por oraciones. La sencilla aseveración siguiente:

(1) Toda oración **P** es verdadera o falsa

no es fácil de enunciar usando variables proposicionales en vez de oracionales. Podríamos comenzar por ensayar la siguiente:

(2) Toda proposición p es verdadera o falsa.

Pero si aplicamos especificación universal a (2), reemplazando '*p*' por 'Jerónimo está muerto', obtenemos una oración sin sentido desde el punto de vista gramatical, a saber, una oración que parece tener dos verbos principales:

(3) Jerónimo es muerto es verdadero o falso.

Como segundo ensayo podríamos dar a (2) una nueva redacción:

(4) Para toda proposición p, '*p*' es verdadera o falsa.

Si no retrocediéramos ante la sustitución de '*p*' cuando está entre comillas, podría obtener de (4) la oración con sentido:

(5) 'Jerónimo está muerto' es verdadera o falsa.

[1] Es quizás conveniente hacer la observación de que en la obra de Quine *Methods of Logic* las letras '*p*', '*q*', '*r*' no se usan, sino se mencionan. De aqui que no sean usadas como variables por Quine y lo que se ha dicho anteriormente no implica que tácitamente se haya confinado dentro de proposiciones.

Sin embargo, hay buenas razones para retroceder ante esta sustitución. Considérese, por ejemplo, la oración:

(6) Para toda *p*, la letra '*p*' es la decimoctava letra del alfabeto.

Es de suponerse que esta oración sería generalmente considerada como verdadera y el cuantificador 'para toda *p*' sería visto como redundante e incapaz de ligar la incidencia de la '*p*' entrecomillada. Si consideramos a (6) como verdadera y al cuantificador como de hecho ligando la incidencia de la '*p*' entrecomillada, obtenemos mediante la sustitución usada para obtener (5) de (4) el falso enunciado:

(7) 'Jerónimo está muerto' es la decimoctava letra del alfabeto.

La regla que cierra el paso a (7) puede resumirse como sigue:

(II) *Los cuantificadores que están fuera de las comillas no pueden ligar variables que figuran dentro de las comillas.*

La Regla (II), como la (I), puede hacerse a un lado sólo por razones muy profundas.

Sin pretender haber agotado todas las alternativas en la transcripción de (1) a una oración que use variables proposicionales, creo que podemos con derecho postular que las alternativas estudiadas constituyen un alegato que presenta a las variables oracionales como opuestas a las oracionales.

EJERCICIOS

1. En cada una de las siguientes oraciones decídase si las letras '*A*', '*B*' y '*C*' deben ser variables oracionales o proposicionales.

(a) Para toda *A* y *B*, *A* implica la disminución de *A* y *B*.
(b) Para toda *A*, *B* y *C*, si *A* sólo si *B* y *B* sólo si *C*, entonces *A* sólo si *C*.
(c) Para toda *A* y *B* la implicación cuyo antecedente es la conjunción de *A* y *B* y cuyo consecuente es la disyunción de *A* y *B*, es una tautología.

2. Señale la confusión entre variables proposicionales y oracionales en la frase siguiente:

Si *p* implica a *q* y *q* implica a *p*, entonces *p* si y sólo si *q*.

§6.3. Yuxtaposición de nombres. En el apartado precedente, defendimos la selección de variables oracionales. Pero no mencionamos cierta dificultad que se suscita en su uso. Considérese la oración:

(1) Para todas las oraciones **P** y **Q**, $\mathbf{P} \to (\mathbf{P} \vee \mathbf{Q})$ es verdadera.

Aplicando la especificación universal a (1) y representando las variables oracionales por los nombres de estas oraciones, obtenemos:

(2) 'Jerónimo está muerto' → ('Jerónimo está muerto' ∨ 'Caballo Loco era dos caras') es verdadera.

El problema consiste en que el sujeto de (2) no es una oración; en efecto, tal como está, (2) carece de significado, pues no tiene un adecuado sujeto gramatical. Antes de la expresión 'es verdadera' esperábamos encontrar el nombre de una oración, pero en lugar de esto vemos que 'es verdadera' está precedida de nombres de oraciones, signos de conectivos oracionales y paréntesis.

Para remediar esta situación necesitamos introducir dos convenciones que son fundamentales por lo que hace a los significados adoptados en este libro.

Convención (I), *Los siguientes signos lógicos se usan como nombres de sí mismos:* –, &, ∨, →, ↔, (,), ∃, =.

La Convención (I) patrocina el uso de identidades de aspecto tan inusitado como:

& = '&',

→ = '→'.

Usando (I), podemos reemplazar (2) por (3):

(3) 'Jerónimo está muerto' '→' '(' 'Jerónimo está muerto' '∨' 'Caballo Loco era dos caras' ')' es verdadera.

La segunda convención permite la conversión de (3) en una oración normal.

Convención (II). *Principio de yuxtaposición. Si* $\mathbf{N}_1$ *es el nombre de la expresión* $\mathbf{E}_1$, *y* $\mathbf{N}_2$ *es el nombre de la expresión* $\mathbf{E}_2$, *entonces la expresión consistente en* $\mathbf{N}_1$ *seguida inmediatamente de* $\mathbf{N}_2$, *es el nombre de la expresión consistente en* $\mathbf{E}_1$ *seguida de* $\mathbf{E}_2$.

Las siguientes identidades ejemplifican el principio:

(4)
'Fx' '→' = 'Fx →',
'–' '(' 'Jerónimo está muerto' ')' = '–(' 'Jerónimo está muerto)'
= '–(Jerónimo está muerto)'

O usando (I), podemos escribir en vez de (4):

'*Fx*' → = '*Fx* →'

–('Jerónimo está muerto') = '–(Jerónimo está muerto)'.

Usando (II), podemos rectificar (3) para obtener la oración verdadera:

(5) 'Jerónimo está muerto → (Jerónimo está muerto ∨ Caballo Loco era dos caras)' es verdadera.

La extensión de (II) a la yuxtaposición de más de dos nombres es evidente y ha sido usada en los ejemplos anteriores.[3]

Las Convenciones (I) y (II) han sido tácitamente usadas a través de los capítulos anteriores.[4] Su uso no ha sido restringido a variables oracionales; también se requiere que los enunciados relativos a inferencias con cuantificadores de los Caps. 4 y 5 tengan un sentido exacto y literal. Por ejemplo, considérese el enunciado de la regla de generalización universal:

> De **S** podemos deducir (**v**)**S**, siempre que **v** no esté imprecisa en **S** y que aparezca como subíndice en **S**.

Aquí '**S**' es una variable que reemplazamos por el nombre de una fórmula, esto es, '**S**' es una variable oracional (extendiendo el concepto de oración al más general de fórmula). La letra '**v**', por otra parte, es una variable que reemplazamos por nombres de variables individuales tales como 'x' y 'y'. Así, como ejemplo de *UG* tenemos:

(6) De '$(x + x > x)$' podemos deducir ('x')'$(x + x > x)$'.

Usando (I) y (II) rectificamos (6) y la convertimos en el enunciado legítimo:

(7) De '$x + x > x$' podemos deducir '$(x)(x + x > x)$'.

Ciertas inconsistencias o extensiones de la significación de poca monta, que no están cubiertas por (I) y (II), han sido permitidas en los capítulos anteriores. Por medio de los ejercicios que se dan a

[3] En el lenguaje matemático se usa la yuxtaposición de nombres de oraciones para denotar una operación binaria sobre las expresiones, justamente como la yuxtaposición de variables numéricas se usa para denotar la operación binaria de multiplicación. Algunas veces es conveniente contar con un símbolo de multiplicación y, similarmente, podríamos, si lo deseáramos, introducir el símbolo de Tarski '⌢' de concatenación para denotar la operación apropiada de sobrepares de expresiones. Podríamos entonces escribir, por ejemplo,

'*Fx*'⌢'→' = '*Fx* →'

[4] Hemos también aplicado la convención de imprimir en negrillas todas las variables que son reemplazadas por nombres de expresiones.

continuación se dirige explícitamente la atención hacia algunas de las violaciones. Sobre este particular es quizás necesario decir que en los Caps. 4 y 5 no se usaron ni variables reemplazables por predicados ni variables reemplazables por nombres de predicados. Cuando se necesitan esas variables es necesario hacer una selección similar a la que se hace entre variables oracionales y proposicionales. Se acostumbra decir que una variable que es reemplazada por un predicado toma como valores *funciones proposicionales*, esto es, los predicados denotan funciones proposicionales; una variable que es reemplazada por un nombre de un predicado toma como sus valores *funciones oracionales*, pero las funciones oracionales son, por supuesto, nada más que predicados. Muchos filósofos encuentran aún más difícil creer que existen entidades como las funciones proposicionales, que creer que existen proposiciones. En este libro no se ha tomado una decisión sobre las proposiciones ni sobre las funciones proposicionales. Las letras '*F*', '*G*', '*H*' y similares usadas en los Caps. 4 y 5 no son variables, sino constantes-predicados. Si deseamos ser explícitos sobre su significado, debemos suministrar sus equivalencias:

$$Hx \leftrightarrow x \text{ es humano},$$

y así, sucesivamente.

EJERCICIOS

1. ¿Qué extensión de la Convención (I) se requiere exactamente para cubrir el uso de las oraciones variables de las fórmulas desplegadas del §1.3, que dan **P** solamente si **Q**, etc., como equivalentes de 'si... entonces...'?

2. ¿Qué diferencia hay, por lo que hace al uso y la mención de las fórmulas, entre la forma en que están desarrolladas las deducciones del Cap. 2 y las del 4?

3. Puesto que la yuxtaposición de nombres denota una operación binaria sobre las expresiones, podemos preguntar:

   (a) ¿Es asociativa la operación denotada por la yuxtaposición de nombres?

   (b) ¿Satisface esta operación las leyes de cancelación a mano derecha y a mano izquierda?
   (Para la explicación de esta terminología, véase la discusión y ejercicios sobre teoría de grupos en el capítulo anterior).

   (c) ¿Satisface esta operación el segundo y el tercer axioma sobre grupos introducidos en el Cap. 5?

CAPITULO 7

TRANSICION DE LAS DEMOSTRACIONES FORMALES A LAS INFORMALES

§7.1. Consideraciones generales. No se acostumbra en matemáticas ni en las ciencias empíricas presentar la deducción de las conclusiones de un modo tan formal como el desarrollado en los Caps. 4 y 5. Por otra parte, los estudiantes que comienzan el estudio de las ciencias quedan a menudo desconcertados por los criterios que rigen la aceptación y el rechazo de las demostraciones matemáticas propuestas, cuando éstas son presentadas de manera informal. Por ejemplo, si en una demostración informal un estudiante comete un error al sustituir variables es a menudo difícil convencerlo de que está equivocado, a menos de que se cuente con un bien definido equipo de pruebas formales para dejar aclarado el error.

El propósito de este capítulo es el de ayudar a la transición de las deducciones formales y a las demostraciones informales en el sentido matemático usual. Se experimentará sin duda cierta incomodidad al pasar del dominio neto y preciso de la lógica formal, donde hay una regla para cada cosa, al mundo más complicado y menos preciso de las matemáticas ordinarias y de las ciencias empíricas.

En una demostración informal sólo se enuncia del argumento lo que baste para que la persona familiarizada con el tema siga la línea de pensamiento con un grado relativamente alto de facilidad y comodidad. Se supone que resulta transparente llenar en forma intuitiva las lagunas lógicas de la demostración. En muchos respectos las normas de inteligibilidad de las demostraciones informales son similares a las de una conversación informal. Así, si alguien me pregunta si iré hoy a casa manejando mi coche, supóngase que contesto, "No, porque mi automóvil está siendo reparado en la estación de servicio". No procedo a enunciar todas las demás premisas obvias que, juntamente con la aserción de que mi carro está en la estación

de servicio, implicarían lógicamente la conclusión de que no iré hoy a casa manejando mi automóvil. De modo análogo, al dar una demostración informal, tratamos de efectuar los pasos esenciales y no evidentes y omitir las inferencias triviales y de rutina. Sin embargo, es un lugar común de exacta filosofía que los conceptos de ser esencial, ser desconocido o ser trivial no son precisos y no se hacen precisos fácilmente. La misma vaguedad de los criterios que rigen las demostraciones informales es una justificación *prima facie* de la necesidad de una definición precisa de demostración formal. En caso de controversia o duda sobre la validez de una demostración informal, es sumamente útil tener a mano una norma exacta y clara a la cual apelar.

Nuestro método para resolver este problema de la transición consistirá en introducir cierto número de axiomas que expresen hechos básicos sobre los números, y a continuación deducir tanto formal como informalmente algunas conclusiones intuitivamente bien conocidas. Por tres razones se han escogido inferencias lógicas relacionadas con la aritmética, para ilustrar los métodos de demostración informal. En primer lugar, el material es intuitivamente bien conocido de todos desde la niñez. En segundo lugar, los ejemplos de inferencias en este dominio tienen un genuino contenido matemático, en el sentido de que las inferencias no son intuitivamente triviales. Muchos de los ejercicios del Cap. 4 no tienen contenido matemático en este sentido, porque a menudo es intuitivamente obvio que las conclusiones se siguen lógicamente de las premisas dadas en los ejercicios. En comparación, no es de ninguna manera obvio que todas las propiedades aritméticas de la adición, la multiplicación, y la sustracción puedan deducirse de los quince axiomas básicos presentados en el apartado siguiente. En tercer lugar, sólo se necesitan unos cuantos más axiomas para sumarse a los que se deben introducir para proporcionar una base sistemática para el desarrollo lógico de los cálculos diferenciales e integrales, la rama de las matemáticas que ha sido hasta hoy, la más importante en las ciencias empíricas.

§7.2. Axiomas básicos sobre números. Esperamos que sean inmediatamente reconocidos cada uno de los siguientes axiomas como una conocida verdad de aritmética. Además del aparato general de la lógica, los axiomas usan solamente las constantes individuales '0' y '1', los bien conocidos símbolos de operación '+' y '·', que denotan las operaciones de adición y multiplicación, respectivamente, y el símbolo de relación '<', que denota la relación *menor que*.[1]

[1] Los axiomas son, esencialmente, los dados por Tarski en su *Introducción a la Lógica*.

(1) $(x)(y)(x + y = y + x)$
(2) $(x)(y)(x \cdot y = y \cdot x)$
(3) $(x)(y)(z)[(x + y) + z = x + (y + z)]$
(4) $(x)(y)(z)[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$
(5) $(x)(y)(z)[x \cdot (y + z) = (x \cdot y) + (x \cdot z)]$
(6) $(x)(x + 0 = x)$
(7) $(x)(x \cdot 1 = x)$
(8) $(x)(\exists y)(x + y = 0)$
(9) $(x)(y)[y \neq 0 \rightarrow (\exists z)(x = y \cdot z)]$
(10) $(x)(y)[x < y \rightarrow -(y < x)]$
(11) $(x)(y)(z)[(x < y \,\&\, y < z) \rightarrow x < z]$
(12) $(x)(y)[x \neq y \rightarrow (x < y \vee y < x)]$
(13) $(x)(y)(z)[y < z \rightarrow (x + y < x + z)]$
(14) $(x)(y)(z)[(0 < x \,\&\, y < z) \rightarrow x \cdot y < x \cdot z]$
(15) $0 \neq 1$

Puesto que muchos de estos axiomas tienen descripciones verbales bien conocidas, es conveniente examinar cada uno. Los axiomas 1 y 2 aseveran que la adición y la multiplicación son ambas operaciones *conmutativas*, esto es, si sumamos o multiplicamos dos números, entonces no importa en qué orden se tomen. Por ejemplo $2 + 3 = 3 + 2$, y $4 \cdot 3 = 3 \cdot 4$. Los axiomas 3 y 4 aseveran que la adición y la multiplicación son ambas operaciones *asociativas*, esto es, el resultado de sumar o multiplicar tres números es independiente de la forma en que se agrupen. El Axioma 5 dice que la multiplicación es *distributiva* con respecto de la adición; esta propiedad distributiva es lo que permite "sacar como factor común" un factor que aparece en todos los términos de una suma. El Axioma 6 dice que la adición de cero a cualquier número simplemente da de nuevo ese número. Más técnicamente, el contenido de este axioma es que cero es un *elemento de identidad a mano derecha para la adición*. (En general, un elemento e es un elemento de identidad a mano derecha para una operación $\circ$ si para toda x

$$x \circ e = x.$$

Un elemento de identidad a mano izquierda se define de manera símilar. Un elemento que es tanto un elemento de identidad a mano izquierda como a mano derecha, se llama simplemente *elemento de identidad*). El Axioma 7 dice que 1 es un elemento de identidad a mano derecha para la multiplicación. Dentro de poco probaremos que tanto 0 como 1 son, en realidad, elementos de identidad únicos. El Axioma 8 dice que por cada número x existe un número y tal que $x + y = 0$. Posteriormente demostraremos (lo que ya sabemos intuitivamente) que de nuestros axiomas se sigue que y es única. El

elemento y es el *elemento inverso* de x (con respecto de la adición). El Axioma 9 dice que si x es un número y y es un número no igual a 0, entonces podemos encontrar un número z tal que $y \cdot z = x$. Como en el caso del Axioma 8, no es difícil demostrar que esta z es única. El contenido intuitivo de este axioma es que la división es siempre posible, excepto entre cero. Los Axiomas 10 a 12 afirman propiedades de ordenacion bien conocidas de la relación $<$. El Axioma 13 asevera que la adición del mismo número a dos números dados preserva la desigualdad entre los números dados, y el Axioma 14 enuncia que multiplicar los dos miembros de una igualdad por un número positivo no altera la igualdad. El axioma final, Axioma 15, indica que '0' y '1' designan elementos distintos.

Los axiomas que hemos estado discutiendo son satisfechos por el conjunto de todos los números racionales (un número racional es un número igual a la razón de dos enteros) y también por el conjunto de todos los números reales. Por el momento, podemos caracterizar el conjunto de los números reales como el conjunto de todos los decimales ilimitados. Por ejemplo, $\sqrt{2}$ es un número real, pero no racional, puesto que no es igual a la razón de dos enteros; por otra parte, todo número racional es un número real. Por ejemplo, el número racional $4/3$ es idéntico al racional ilimitado 1.3333 ... En otras palabras, podemos escoger como nuestro dominio de individuos, o el conjunto de números racionales o el conjunto más amplio, esto es, más inclusivo, de números reales. Como veremos, los Axiomas 1 a 15 son adecuados para desarrollar las propiedades aritméticas de las operaciones fundamentales de adición, multiplicación y sustracción.

EJERCICIOS

1. ¿Cuál de los Axiomas de 1 a 15 no son satisfechos por
   (a) los enteros positivos,
   (b) los enteros y
   (c) los números racionales no negativos?
2. Una operación $\circ$ es *conmutativa* si para toda x y y

$$x \circ y = y \circ x.$$

Los axiomas 1 y 2 dicen que la adición y multiplicación son operaciones conmutativas.

   (a) ¿Es conmutativa la sustracción?
   (b) ¿Es conmutativa la división?

3. Una operación $\circ$ es *asociativa* si para cada x, y y z,

$$(x \circ y) \circ z = x \circ (y \circ z).$$

Los axiomas 3 y 4 dicen que la adición y multiplicación son operaciones asociativas.

(a) ¿Es asociativa la sustracción? Si no lo es, dése un contraejemplo explícito.
(b) ¿Es asociativa la división? Si no lo es, dése un contraejemplo explícito.

4. Una operación $\star$ es *disrtributiva* con respecto de una operación $\circ$ si para toda x, y y z,

$$x \star (y \circ z) = (x \star y) \circ (x \star z).$$

El Axioma 5 dice que la multiplicación es distributiva con respecto de la adición. Cuando es negativa la respuesta a cualquiera de las cuestiones siguientes, dése un contraejemplo explícito.

(a) ¿Es distributiva la adición con respecto de la multiplicación?
(b) ¿Es distributiva la sustracción con respecto de la multiplicación?
(c) ¿Es distributiva la multiplicación con respecto de la sustracción?
(d) ¿Es distributiva la sustracción con respecto de la división?

5. ¿Existe un elemento de identidad a mano izquierda o a mano derecha para la sustracción? .

6. Dé una interpretación que pruebe que el Axioma 5 es independiente de los otros catorce. (Advertencia: no viole la Regla VII del §4.2).

7. Dé una interpretación que demuestre que el Axioma 8 es independiente de los otros catorce. (Advertencia: no viole la Regla VII del §4.2).

8. Dé una interpretación que demuestre que el Axioma 9 es independiente de los otros catorce. (Advertencia: no viole la Regla VII del §4.2).

§7.3. Ejemplos comparativos de deducciones formales y de demostraciones informales. Volvemos ahora al principal asunto de este capítulo: la introducción de las demostraciones informales. Usaremos como nuestras premisas básicas los quince axiomas admitidos en el apartado anterior. Las deducciones formales comenzarán, pues, en la línea (16), siguiendo las reglas establecidas en capítulos anteriores. Nuestra única abreviatura consistirá en hacer múltiples aplicaciones de la regla que rige las identidades. En el Cap. 5 se dio el nombre de teorema de lógica a cualquier fórmula deducible del conjunto vacío de premisas. Llamamos aquí *teorema de aritmética* a cualquier fórmula deducible de nuestros axiomas básicos.

Es importante advertir que hay una enorme diferencia entre la tarea relativamente trivial de reconocer la verdad de los teoremas que enunciaremos y la empresa mucho más difícil de demostrar que son de hecho *consecuencias lógicas* de los quince axiomas enunciados en la sección anterior. El lector no debe desconcertarse por el contenido aparentemente trivial de algunos de los teoremas. En esos casos, es posible que se incline a pensar: "Esto es tan obvio, que resulta tonto pedirme que lo demuestre". Pero si el

lector piensa esto, está confundiendo la evidencia de la verdad del teorema con la evidencia de la demostración de que el teorema es una consecuencia lógica de los axiomas.

Nuestro primer teorema enuncia que cero es el único elemento de identidad a mano derecha de la adición. En la deducción formal se omiten las referencias a *US*, puesto que se indican las sustituciones que de hecho se hacen y la notación de sustitución significa una aplicación de *US*.

TEOREMA 1. $(x)[(y)(y + x = y) \rightarrow x = 0]$

DEDUCCIÓN FORMAL

{16}	(16) $(y)(y + x = y)$	x P
{16}	(17) $0 + x = 0$	x 16 $0/y$
{6}	(18) $x + 0 = x$	6 x/x
{1}	(19) $x + 0 = 0 + x$	1 x/x, $0/y$
{1, 6, 16}	(20) $x = 0$	x 17, 18, 19 I
{1, 6}	(21) $(y)(y + x = y) \rightarrow x = 0$	16, 20 C.P.
{1, 6}	(22) $(x)[(y)(y + x = y) \rightarrow x = 0]$	21 *UG*

DEMOSTRACIÓN INFORMAL. Por hipótesis del teorema, para toda y, $y + x = y$. De aquí que, (sustituyendo 'y' por '0'), obtenemos:

$$0 + x = 0$$

Pero el axioma 6 asevera que $x + 0 = x$. Puesto que, en virtud del Axioma 1, la adición es conmutativa, esto es, $x + 0 = 0 + x$, inferimos inmediatamente que $x = 0$, Q.E.D.

Como es de esperarse de lo dicho anteriormente, la demostración informal no menciona explícitamente ninguno de los usos de las reglas lógicas de inferencia; sólo se mencionan por su nombre axiomas específicos. En otras palabras, las reglas formales de derivación se dan por cumplidas. La expresión entre paréntesis de la prueba informal es una concesión a la mayor claridad, que no siempre se hace. Las sustituciones particulares hechas en axiomas o en teoremas previos no se describen, a menos que sean complicadas o nada evidentes. Excepto cuando una prueba informal es excepcionalmente larga, no hay "resumen" en la forma rutinaria que aparece en las líneas (21) y (22) de la deducción formal. La obvia tarea establecida por el Teorema 1 es la de demostrar que '$x = 0$' se sigue de la hipótesis del teorema. Cuando se hace esto, la demostración informal se lleva a cabo satisfactoriamente. El 'Q.E.D.' final, que son las siglas de la expresión *quod erat demonstrandum*, es una forma tradicional de indicar que la demostración ha terminado.

La regla práctica seguida en las demostraciones informales que se han dado en este capítulo y que se recomienda para los ejercicios es:

En una prueba informal indíquese explícitamente CADA *aplicación de un axioma o teorema previo de aritmética. Suprímase ordinariamente la mención de las reglas lógicas de inferencia usadas.*

Puesto que la demostración del segundo teorema, teorema que establece que 1 es el único elemento de identidad a mano derecha para la multiplicación, es muy semejante a la del Teorema 1, la dejamos como ejercicio.

TEOREMA 2. $(x)[(y)(y \cdot x = y) \rightarrow x = 1]$.

El siguiente teorema enuncia la ley de la cancelación para la adición. Su demostración debe compararse con la demostración correspondiente para grupos arbitrarios, del Cap. 5. La diferencia básica estriba en que no hemos definido el concepto de elemento inverso (con respecto de la adición) y por lo mismo no podemos seguir completamente la estrategia de la demostración anterior. En la demostración de la ley de la cancelación para la adición se aplica cómodamente una extensión de la regla que rige las identidades. Designamos a esta extensión con las siglas I2.

REGLA ADICIONAL PARA IDENTIDADES: I2. *Si* $\mathbf{t}_1$, $\mathbf{t}_2$ *y* $\mathbf{t}_3$ *son términos cualesquiera y* $\circ$ *es cualquier símbolo de operación, entonces de* $\mathbf{t}_2 = \mathbf{t}_3$ *podemos deducir* $\mathbf{t}_1 \circ \mathbf{t}_2 = \mathbf{t}_1 \circ \mathbf{t}_3$ *y* $\mathbf{t}_2 \circ \mathbf{t}_1 = \mathbf{t}_3 \circ \mathbf{t}_1$.

DEMOSTRACIÓN: Tenemos como verdad de lógica

(1) $$\mathbf{t}_1 \circ \mathbf{t}_2 = \mathbf{t}_1 \circ \mathbf{t}_2.$$

Suponiendo ahora que es válida la fórmula $\mathbf{t}_2 = \mathbf{t}_3$, aplicamos la regla original para identidades y sustituimos la segunda incidencia de $\mathbf{t}_2$ en (1) por $\mathbf{t}_3$, con lo que obtenemos

$$\mathbf{t}_1 \circ \mathbf{t}_2 = \mathbf{t}_1 \circ \mathbf{t}_3.$$

Mediante un argumento similar obtenemos $\mathbf{t}_2 \circ \mathbf{t}_1 = \mathbf{t}_3 \circ \mathbf{t}_1$.

Como caso especial de esta regla obtenemos por adición y multiplicación como teoremas de lógica:

(1) $$(x)(y)(z)(y = z \rightarrow x + y = x + z),$$

(2) $$(x)(y)(z)(y = z \rightarrow x \cdot y = x \cdot z).$$

No es desusado ver oraciones como (1) y (2) dadas como axiomas

de algún dominio especial de las matemáticas, como la aritmética, pero esta práctica descansa sobre una confusión, pues estas dos oraciones son verdades de lógica y es redundante añadirlas como nuevos axiomas de la materia. El tercer axioma es llamado por razones obvias la *ley de la cancelación* para la adición.

TEOREMA 3. $(x)(y)(z)(x + y = x + z \rightarrow y = z)$.

DEDUCCION FORMAL

{8}	(16) $(\exists y)(x + y = 0)$	8 x/x
{8}	(17) $x + \alpha_x = 0$	16 *ES*
{1}	(18) $x + \alpha_x = \alpha_x + x$	1 x/x, α_x/y
{1, 8}	(19) $\alpha_x + x = 0$	17, 18 I
{1}	(20) $0 + y = y + 0$	1 $0/x$, y/y
{6}	(21) $y + 0 = y$	6 y/x
{1, 6}	(22) $0 + y = y$	20, 21 I
{1, 6, 8}	(23) $(\alpha_x + x) + y = y$	19, 22 I
{3}	(24) $(\alpha_x + x) + y = \alpha_x + (x + y)$	3 α_x/x, x/y, y/z
{6}	(25) $z + 0 = z$	6 z/x
{1}	(26) $0 + z = z + 0$	1 $0/x$, z/y
{1, 6, 8}	(27) $(\alpha_x + x) + z = z$	19, 25, 26 I
{3}	(28) $(\alpha_x + x) + z = \alpha_x + (x + z)$	3 α_x/x; x/y, z/z
{29}	(29) $x + y = x + z$	x, y, z P
{29}	(30) $\alpha_x + (x + y) = \alpha_x + (x + z)$	x, y, z 29 I2
{1, 3, 6, 8, 29}	(31) $y = z$	y, z, 23, 24, 27, 28, 30 I
{1, 3, 6, 8}	(32) $x + y = x + z \rightarrow y = z$	29, 31 C.P.
{1, 3, 6, 8}	(33) $(x)(y)(z)(x + y = x + z \rightarrow y = z)$	32 *UG*

DEMOSTRACIÓN INFORMAL. Según el Axioma 8 hay un número, digamos u, tal que $x + u = 0$, y de aquí que, por la ley conmutativa para la adición:

(1) $$u + x = 0.$$

Aplicando el Axioma 6 y las leyes conmutativa y asociativa para la adición, obtenemos la siguiente serie de identidades:

$$\begin{aligned} y &= y + 0 && \text{(Axioma 6)} \\ &= 0 + y && \text{(Ley conmutativa)} \end{aligned}$$

$= (u + x) + y$	(Según (1))
$= u + (x + y)$	(Ley asociativa)
$= u + (x + z)$	(Hipótesis de teorema)
$= (u + x) + z$	(Ley asociativa)
$= 0 + z$	(Según (1))
$= z + 0$	(Ley conmutativa)
$= z$	(Axioma 6) Q.E.D.

La notación a la derecha de la prueba informal debe explicarse por sí misma; por ejemplo, la primera identidad se sigue del Axioma 6 y la segunda identidad, $y = 0 + y$ se sigue de la primera por la ley conmutativa de la adición (y la regla lógica que rige las identidades). En varios respectos la estructura de la demostración informal es más clara que la deducción formal; en particular, la línea principal de la argumentación se presenta sin embarazosos requisitos, como las líneas (18), (22), (25) a (28) de la deducción formal, para no mencionar el redundante resumen de (32) y (33). La notación usada a la derecha ejemplifica una tendencia general de las demostraciones informales: las propiedades tipo, como la conmutatividad de la adición, son designadas por su nombre, más bien que por referencia al axioma o teorema correspondiente. Nótese que en la demostración informal 'u' desempeña el papel de nombre ambiguo; corresponde a 'α_x' de la deducción formal. La expresión, 'digamos u', al principio de la demostración informal, la usan a menudo los matemáticos para indicar que una letra que generalmente se usa como variable, está siendo usada para fines inmediatos, como nombre ambiguo.

No se trata de dar la impresión de que hay exactamente un estilo correcto que debe emplearse en la formulación de demostraciones informales, en el sentido en que un soneto consta de catorce endecasílabos, agrupados en dos cuartetos y dos tercetos. Algunos matemáticos criticarían la prueba informal del Teorema 3 por no tener redacción suficiente, pues la indicación de cada paso de la prueba no se expresa en una oración completa y bien formada, sino se usa más bien una notación bárbara de frases incompletas entre paréntesis. Mayor número de matemáticos aún criticaría la indebida longitud de las dos demostraciones informales dadas hasta ahora, pero esta última crítica podría quizás ser rechazada basándose en que cuando se principia a formular demostraciones informales es mejor ser claro y un poco prolijo que breve pero confuso.

El cuarto teorema enuncia el hecho familiar de que la multiplicación de cualquier número por cero da como resultado cero.

TEOREMA 4. $(x)(x \cdot 0 = 0)$.

DEDUCCION FORMAL

{6}	(16) $x + 0 = x$	6 x/x
{6}	(17) $(x \cdot x) + 0 = x \cdot x$	6 $(x \cdot x)/x$
{6}	(18) $(x \cdot x) + 0$ $= x \cdot (x + 0)$	16, 17 I
{5}	(19) $x \cdot (x + 0)$ $= (x \cdot x) + (x \cdot 0)$	5 x/x, x/y, $0/z$
{5, 6}	(20) $(x \cdot x) + (x \cdot 0)$ $= (x \cdot x) + 0$	18, 19 I
{1, 3, 6, 8}	(21) $(x \cdot x) + (x \cdot 0)$ $= (x \cdot x) + 0$ $\rightarrow (x \cdot 0) = 0$	Teo. 3 $(x \cdot x)/x$, $(x \cdot 0)/y$, $0/z$
{1, 3, 5, 6, 8}	(22) $x \cdot 0 = 0$	20, 21 T
{1, 3, 5, 6, 8}	(23) $(x)(x \cdot 0 = 0)$	22 *UG*

DEMOSTRACIÓN INFORMAL 1. Tenemos las siguientes identidades:

$$\begin{aligned} (x \cdot x) + (x \cdot 0) &= x \cdot (x + 0) && \text{(Ley distributiva)} \\ &= x \cdot x && \text{(Axioma 6)} \\ &= (x \cdot x) + 0 && \text{(Axioma 6 de nuevo).} \end{aligned}$$

De aquí, poniendo '$x \cdot x$' en lugar de 'x', '$x \cdot 0$' en lugar de 'y' y '0' en lugar de 'z' en el Teorema 3, obtenemos inmediatamente el resultado $x \cdot 0 = 0$. Q.E.D.

Puesto que muchos matemáticos considerarían como inelegante la lista de las sustituciones hechas en el Teorema 3, podemos formular de nuevo esta demostración informal y eliminar este uso. Nótese que el Teorema 3 se cita ahora por el nombre de la propiedad que expresa.

DEMOSTRACIÓN INFORMAL 2. Puesto que por la ley de distribución y el Axioma 6,

$$(x \cdot x) + (x \cdot 0) = x \cdot (x + 0) = x \cdot x = (x \cdot x) + 0,$$

podemos usar la ley de la cancelación para la adición a fin de obtener $x \cdot 0 = 0$. Q.E.D.

Las observaciones sobre pruebas informales contenidas en este apartado han sido casi enteramente de carácter estilístico. Los problemas estilísticos son sin duda serios para los que empiezan a adquirir una firme noción de la prueba matemática. Pero quizás son más importantes los problemas que se suscitan de los esfuerzos por

desarrollar una *estrategia* eficiente para encontrar demostraciones. No existen, por supuesto, métodos precisos para desarrollar una buena estrategia, pero trataremos de dar sobre la marcha algunas útiles sugestiones. En la prueba del Teorema 4, por ejemplo, lo natural es comenzar por observar determinadamente los axiomas y escoger los que contienen cero. El Axioma 6 parece el más promisor. En vista de los métodos aplicados para demostrar el teorema precedente, el paso siguiente sería el de considerar si podría hacerse alguna aplicación de las leyes, conmutativa, asociativa o distributiva. Cuando se haya escogido la ley distributiva, habrá comenzado a tomar forma una demostración. Estas observaciones son, por decirlo así, *a posteriori*; es un buen elemento de una buena estrategia estar en disposición de ensayar cierto número de procedimientos diferentes. Si el procedimiento que se acaba de describir no da resultado, entonces debe tratarse alguna otra cosa que incluya quizás algunos de los axiomas aún no aplicados.

Ciertas estructuras generales vuelven a presentarse un tanto a menudo en las demostraciones, y es conveniente estar al tanto de ellas. En las demostraciones de los Teoremas 3 y 4 se tiene un ejemplo de esas estructuras; las pruebas consisten fundamentalmente en una sarta de identidades. Se tiene un ejemplo de una segunda clase de estructura en la demostración del Teorema 6 del § 7.5: la prueba consiste principalmente en una serie de implicaciones. Una tercera clase de estructura es suministrada por el familiar método de la prueba indirecta, que primero usamos en este capítulo en la demostración del Teorema 14 del § 7.5.

El Teorema 4 es el último teorema para el cual damos una deducción formal. Consecuentemente, en el enunciado subsecuente de teoremas (o de axiomas), omitiremos usualmente los cuantificadores universales iniciales cuyos alcances se extiendan al resto de la fórmula, lo que constituye una práctica usual en matemáticas. En las deducciones formales no se tiene una regla que permita sustituir las variables libres como ocurre en los teoremas, cuando se omiten los cuantificadores universales, pero necesitamos aplicar estos teoremas a situaciones que implican variables distintas de las que aparecen en las formulaciones iniciales de los teoremas. Como podía esperarse, se acostumbra en las demostraciones informales sustituir las variables libres en los axiomas y teoremas cada vez que sea necesario. Por otra parte, si en la prueba de un teorema basado en los axiomas, introducimos el antecedente del teorema como una premisa (véase, por ejemplo, el Teorema 3) tratamos las variables

libres de esta premisa como imprecisas, y no las sustituimos.[2] La razón intuitiva de esto es obvia. Supóngase que el antecedente contiene las variables libres 'x', 'y', y 'z'. Está claro que no bastará con probar el teorema para el caso especial de $x = y = z$; esto es, no estableceríamos el teorema en toda su generalidad si sustituyéramos 'y' y 'z' por 'x'. La sección siguiente está dedicada al estudio de esos métodos falaces. También se hace una ulterior aclaración del uso de variables libres.

EJERCICIOS

1. Demuéstrese formal e informalmente el Teorema 2.

2. Dense deducciones formales y las correspondientes demostraciones informales de lo siguiente:

(a) $(x)-(x < x)$
(b) $(x)(y)(x = y \rightarrow -(x < y \vee y < x))$
(c) $(x)(y)(x < y \rightarrow -(x = y \vee y < x))$
(d) $(x)(y)(z)(x + y < x + z \rightarrow y < z)$

3. Dada la premisa adicional

$$(16) \qquad (x)(y)(x \leq y \leftrightarrow (x = y \vee x < y)),$$

pruébese formal e informalmente que

(a) $(x)(x \leq x)$
(b) $(x)(y)[(x \leq y \,\&\, y \leq x) \rightarrow x = y]$
(c) $(x)(y)(z)[(x \leq y \,\&\, y \leq z) \rightarrow x \leq z]$
(d) $(x)(y)(x \leq y \vee y \leq x)$

4. Dada la premisa adicional

$$(16) \qquad 2 = 1 + 1$$
$$(17) \qquad 3 = 2 + 1$$
$$(18) \qquad 4 = 3 + 1,$$

pruébese formal e informalmente que $2.2 = 4$.

§7.4. Ejemplos de demostraciones informales falaces. En los Caps 4 y 5 se presentó cierto número de deducciones formales falaces para justificar varias restricciones de las reglas formales de inferencia, y para ejemplificar algunos de los errores más comunes. Es propósito de este apartado el de señalar algunas de las fallas cometidas en las demostraciones informales. Como era de esperarse, los errores en las demostraciones informales tienen sus análogos en las deducciones formales. Infortunadamente, sin embargo, puede uno sorprenderse cometiendo errores en las demostraciones in-

[2] Para un enunciado exacto de una regla que rige la sustitución de las variables libres, véase el Ejercicio 6, § 5.3.

formales que no se cometerían en las deducciones formales. Algunas de las razones de esa discrepancia son difíciles de determinar. Una vez que se abandona la adhesión explícita y exacta a las reglas formales de inferencia hay cierta tendencia natural a pensar que ahora todo se permite. Hay también cierto desconcierto sobre lo que ahora debe considerarse como el correcto enunciado de las partes de un razonamiento y lo que no debe considerarse como tal. Como dirían algunos psicólogos, a fin de hacer la transición de las deducciones formales a las informales, debemos desarrollar cierta tolerancia de la ambigüedad.

Hay modos positivos de considerar esta ambigüedad. La principal regla aproximada es ver una demostración informal como la abreviatura de una deducción formal. No debe haber un solo paso de una demostración informal que no pueda ser desmenuzado o analizado en un razonable número de pasos formales. No puede recalcarse bastante que cada vez que se tenga una duda sobre un paso en particular de una demostración informal, lo que debe hacerse es tratar de desarrollar ese paso y convertirlo en un fragmento de una deducción formal. Si no se puede hacer este desarrollo, y aclararse completamente por la vía formal la validez de ese paso, entonces hay buenas razones para dudar de su validez lógica.

La siguiente demostración ilustra el que probablemente es el más frecuente tipo de error que se comete en las demostraciones informales.

TEOREMA. *Para toda x, y y z, si $x + y < x + z$ entonces $y < z$.*

DEMOSTRACIÓN FALAZ. Por la hipótesis del teorema,

(1) $$x + y < x + z.$$

Poniendo '0' en lugar de 'x' en (1), se obtiene

(2) $$0 + y < 0 + z.$$

De los Axiomas 1 y 6 se tiene:

(3) $$0 + y = y + 0 = y$$

y.

(4) $$0 + z = z + 0 = z.$$

De aquí, usando las identidades (3) y (4), inferimos de (2) que

$$y < z. \qquad \text{Q.E.D.}$$

El error se comete al inferir (2) de (1). No es posible sustituir las variables 'x', 'y' y 'z' de (1). ¿Por qué? Porque desde un punto de vista lógico estas variables están imprecisas, de manera que no es

posible generalizar (1) universalmente y a continuación aplicarle especificación universal para hacer la sustitución apropiada. Así pues, si enfocamos nuestro microscopio lógico sobre la inferencia propuesta de (2) a (1), obtenemos la siguiente deducción falaz:

{1}	(1) $x + y < x + z$	x, y, z P
{1}	(2) $(x)(x + y < x + z)$	1 *UG* (falazmente)
{1}	(3) $0 + y < 0 + z$	2 *US*

Puesto que el error cometido en esta prueba es tan común, no estará de más una ulterior reflexión sobre el mismo. Generalmente, cuando deseamos sustituir las variables libres en una prueba informal, no pensamos en generalizar universalmente y a continuación especificar universalmente. Como se hizo notar en el apartado anterior, puede aplicarse una regla deducida que permita la sustitución directa. Esta regla se da en el ejercicio 6 del § 5.3 y dice lo siguiente:

> *De* **S(v)** *podemos deducir* **S(t)**, *siempre que* (i)**v** *no esté imprecisa,* (ii)**v** *no sea un subíndice y* (iii) *ninguna incidencia libre de* **v** *quede dentro del alcance de un cuantificador que use una variable de* **t**.

Puesto que la regla simplemente telescopia una aplicación de *UG* seguida de una aplicación de *US* en un solo paso, las restricciones impuestas a la misma son justamente las restricciones impuestas a *UG* y *US*. Para uso en pruebas informales podemos dar a la regla una nueva redacción que tome en cuenta cada caso de sustitución que se necesite.

> Regla Informal para Sustituir Variables Libres. (i) *Se pueden colocar términos en lugar de variables que figuren en axiomas o teoremas demostrados anteriormente.* (ii) *Cualquier caso no previsto por* (i) *debe ser remitido a la regla formal; en particular no es posible sustituir variables libres de la hipótesis del teorema que se trata de probar.*

Como se hizo notar en el apartado anterior, de aquí en adelante se omitirán los cuantificadores universales que estén al comienzo de los teoremas. Se configura especialmente la regla informal anterior para permitir la sustitución en estos teoremas. Por ejemplo, siguiendo la convención que está ahora en vigor con respecto de los cuantificadores universales, enunciaríamos el Teorema 3 como:

Teorema 3. *Si* $x + y = x + z$, *entonces* $y = z$.

Se nos permitiría entonces sustituir las variables 'x', 'y' y 'z' en las

aplicaciones del Teorema 3, tal como ha quedado sancionado por la anterior regla informal.

En el ejemplo anterior de demostración falaz consideramos un teorema genuino, para el cual existe una demostración correcta (véase el Ejercicio 2d del apartado anterior). El ejemplo siguiente es un teorema falaz para el cual no existe demostración correcta.

TEOREMA FALAZ. $x = 1$.

(Así, este teorema falaz asevera que todo número es igual a uno. La variable 'x' ha sido dejada libre, de acuerdo con la convención que se acaba de establecer).

DEMOSTRACIÓN FALAZ. Sea x un número arbitrario. Por el Axioma 2 y el Teorema 4, se tiene:

(1) $$0 \cdot x = x \cdot 0 = 0.$$

De aquí, haciendo $y = 0$ en el Teorema 2, inferimos de (1) que

$$x = 1. \qquad \text{Q.E.D.}$$

En términos lógicos, el error cometido en esta prueba consiste en aplicar la especificación universal al cuantificador que usa 'y' en el Teorema 2, pues el alcance de este cuantificador no es toda la fórmula. Se hizo una enérgica advertencia contra esta clase de error en el Cap 4, pero es precisamente el tipo de error que se cometería en una demostración informal y, en cambio, no se cometería en una deducción formal. A fin de hacer uso del Teorema 2, se tendría que establecer que para toda y, no solamente para 0, x tiene la propiedad de que

$$y \cdot x = y.$$

El uso de la expresión 'haciendo $y = 0$ en el Teorema 2, inferimos. . .', no debe hacer que se descuide el análisis de la demostración falaz. Esta expresión no es más que jerga matemática, usada en vez de: aplicando *US* al Teorema 2 mediante la sustitución de 'y' por '0', inferimos. . . Los matemáticos a menudo usan el signo de identidad para indicar una aplicación de *US*. Debe aprenderse a distinguir estos casos de los que representan una genuina aplicación de la regla que rige las identidades. El contexto de aplicación generalmente debe hacer que esta diferenciación resulte un tanto sencilla.

En el ejercicio que sigue se dan otros ejemplos de pruebas informales.

EJERCICIOS

En cada una de las siguientes demostraciones falaces, localícese el error e indíquese qué error lógico se ha cometido. Se dan pruebas falaces de enunciados falsos y verdaderos.

1. TEOREMA 3. *Si* $x + y = x + z$ *entonces* $y = z$.

PRUEBA FALAZ. Por la hipótesis del teorema,

(1) $$x + y = x + z.$$

Haciendo $x = 0$ en (1) se tiene,

(2) $$0 + y = 0 + z.$$

De los axiomas 1 y 6 se tienen las identidades

$$0 + y = y + 0 = y$$

y

$$0 + z = z + 0 = z.$$

De donde, usando estas identidades, se obtiene de (2) que

$$y = z.$$ Q.E.D.

2. TEOREMA FALAZ. *Para toda* x *y* y, *si* $x \cdot y = 1$, *entonces* $x = 1$.

PRUEBA FALAZ. Por hipótesis

(1) $$x \cdot y = 1.$$

Haciendo $y = 1$ se tiene

(2) $$x \cdot 1 = 1.$$

Pero por el Axioma 7

(3) $$x \cdot 1 = x.$$

Se sigue inmediatamente de (2) y (3) que

$$x = 1.$$ Q.E.D.

3. TEOREMA FALAZ. *Si hay una* y *tal que* $x \cdot y = 1$ *entonces* $x = 1$.

PRUEBA FALAZ. Por hipótesis, hay una y tal que

$$x \cdot y = 1.$$

Haciendo $y = 1$, tenemos entonces:

$$x \cdot 1 = 1.$$

Pero por el Axioma 7

$$x \cdot 1 = x.$$

De aquí que

$$x = 1.$$ Q.E.D.

4. TEOREMA FALAZ. *Hay una* x *tal que para toda* y, $x + y = x$.

PRUEBA FALAZ. Por el Axioma 6

$$x + 0 = x.$$

Sea $y = 0$; se tiene, entonces, inmediatamente:

$$x + y = x. \qquad \text{Q.E.D.}$$

§7.5. Otros ejemplos de demostraciones informales. A continuación del desarrollo sistemático de los números racionales y reales, suministraremos una oportunidad de estudiar más pruebas informales. En este apartado definiremos también la operación negativa y la operación de sustracción. Desde un punto de vista formal, una definición es simplemente una premisa adicional. El carácter lógico de las definiciones y los problemas lógicos que generan se examinan pormenorizadamente en el capítulo siguiente. El único punto esencial necesario para este apartado es que una definición que introduce un nuevo símbolo de operación necesita ser precedida por un teorema que garantice que la definición introduce una operación definida unívocamente. El primer teorema de este apartado justifica en este sentido la introducción de la operación negativa. Nótese que la variable 'x' está libre, de acuerdo con la estipulación previamente establecida.

TEOREMA 5. *Hay exactamente una y tal que* $x + y = 0$.

DEMOSTRACIÓN. Por el Axioma 8, hay *cuando menos* una y tal que

(1) $$x + y = 0.$$

Así pues, para probar el teorema necesitamos demostrar que hay *a lo sumo* una y como esa. Sea y' un número tal que

(2) $$x + y' = 0.$$

Entonces

(3) $$x + y = x + y'$$

y por la ley de cancelación

(4) $$y = y'. \qquad \text{Q.E.D.}$$

Esta demostración ejemplifica la forma típica en que subdividimos la demostración de que existe *exactamente* una entidad que satisface cierta condición. Primeramente demostramos que existe cuando menos una, y a continuación que a lo sumo existe una, pues decir que existe exactamente una, equivale a decir que hay cuando menos una y cuando más una. Nótese que de (1) y (2) obtenemos (3), mediante una aplicación tan obvia de la regla que

rige las identidades, que no se da justificación de ninguna clase. También, después de inferir (4) no añadimos la frase redundante 'y en consecuencia existe a lo suma una y', puesto que comenzamos por considerar una y' arbitraria, con $x + y' = 0$. Decir que hay a lo sumo una y tal que $x + y = 0$ es lo mismo que decir que para toda y', *si* $x + y = 0$, entonces $y = y'$. En una deducción normal esto tendría que ser asentado, pero no en una demostración informal.

DEFINICIÓN 1. $-x = y$, *si* y *sólo si* $x + y = 0$.

El *negaivo* de x es, por supuesto, $-x$.

Para aclarar las ideas sobre el uso de definiciones en las demostraciones, necesitamos solamente hacer la observación de que en una deducción formal la Definición 1 sería tratada como la premisa decimosexta:

(16) $$(x)(y)(-x = y \leftrightarrow x + y = 0).$$

A medida que se continúa el desarrollo deductivo de la aritmética (por medio de demostraciones informales), usamos las nuevas operaciones introducidas sin referencia explícita a sus definiciones.

Pasamos ahora a demostrar cinco teoremas que enuncian hechos bien conocidos sobre la operación negativa.

TEOREMA 6. $-(-x) = x$.

DEMOSTRACIÓN. En virtud del Teorema 5, existen números y y z tales que

(1) $$x + y = 0$$

y

(2) $$(-x) + z = 0.$$

De aquí que, usando la Definición 1, tenemos:

(3) $$-x = y$$

(4) $$-(-x) = z.$$

De (2) y (3) obtenemos:

(5) $$y + z = 0.$$

Usando la ley conmutativa para adición, inferimos de (1) y (5) que

$$y + x = y + z,$$

y de aquí, por la ley de cancelación para la adición

$$x = z;$$

esto es, en vista de (4),

$$-(-x) = x. \qquad \text{Q.E.D.}$$

La estrategia usada en esta demostración consiste en tratar a x y su negativo en forma simétrica. Los resultados son a continuación venturosamente combinados en (5). La demostración de este teorema da lugar a un problema que no siempre es fácil de resolver. En el teorema no se ha enunciado ninguna hipótesis, de manera que no existe un supuesto inicial obvio. Además, no hay indicios de una sucesión de identidades, como en el caso del Teorema 8. Consecuentemente, nos vemos obligados a recaer en la definición de la operación negativa y el teorema que justifica la definición (esto es, el Teorema 5). Nuestro problema táctico consiste en encontrar el camino hacia la solución deseada, partiendo de esta tierra firme. La cuestión está aquí en que, en ausencia de las otras dos posibilidades mencionadas, el Teorema 5 y la definición 1 constituyen la base más razonable para la demostración.

TEOREMA 7. $x + (-x) = 0$.

DEMOSTRACIÓN. Es una verdad de lógica que

$$-x = -x,$$

de donde, en virtud de la Definición 1,

$$x + (-x) = 0. \qquad \text{Q.E.D.}$$

Esta definición viola la regla general de no remitir a los principios lógicos. Puesto que implica una aplicación un tanto sutil de la regla que rige las identidades, la violación está justificada. Es conveniente hacer notar que los Axiomas 3 y 6, juntamente con el Teorema 7, son justamente los tres axiomas para un grupo, dados en el Cap 5, donde la adición es la operación de grupo y cero es el elemento de identidad del grupo.

TEOREMA 8. $(-x)\cdot y = -(x\cdot y)$.

DEMOSTRACIÓN. Tenemos las identidades siguientes:

$$\begin{aligned} 0 &= y\cdot 0 && \text{(Teorema 4)} \\ &= y\cdot(x + (-x)) && \text{(Teorema 7)} \\ &= (y\cdot x) + (y\cdot(-x)) && \text{(Ley distributiva)} \\ &= (x\cdot y) + ((-x)\cdot y) && \text{(Ley conmutativa)} \end{aligned}$$

Y, en consecuencia, por la Definición 1,

$$(-x)\cdot y = -(x\cdot y). \qquad \text{Q.E.D.}$$

Como ya se ha hecho notar, la demostración de este teorema sigue un patrón familiar. Dado un teorema que enuncia una identidad, el método más natural es demostrarlo mediante una serie de iden-

tidades. Nótese que la demostración consiste en algo más que las identidades.

Teorema 9. $-y < -x$, *si y sólo si* $x < y$.

Demostración. Primeramente, supongamos que $x < y$. Aplicando el Axioma 13, obtenemos:

$$-x + x < -x + y,$$

y, en consecuencia, por el Teorema 7 y la ley conmutativa para la adición

(1) $$0 < -x + y.$$

Aplicando de nuevo el Axioma 13, obtenemos de (1):

$$-y + 0 < -y + (-x + y),$$

de donde, por el Axioma 6

(2) $$-y < -y + (-x + y).$$

Aplicando ahora las leyes conmutativa y asociativa al segundo miembro de (2), concluimos

$$-y < -x + (-y + y);$$

y, puesto que $-y + y = 0$, obtenemos el resultado deseado:

$$-y < -x.$$

Suponiendo ahora que $-y < -x$, inferimos por un argumento exactamente similar que $x < y$. Q.E.D.

El Teorema 9 es el primero que ha tenido la forma de una equivalencia. Como ya se hizo notar en el Cap 5, primeramente demostramos que un miembro de la equivalencia implica al otro y a continuación establecemos la implicación en la dirección opuesta. En este respecto, a menudo podemos tomarnos libertades en una demostración informal que no serían permitidas en una deducción formal. Así, en la demostración del Teorema 9 demostramos esencialmente sólo una de las dos implicaciones y a continuación simplemente hacemos la observación de que es similar la demostración de la segunda implicación. Naturalmente, si la demostración de la segunda implicación implicara un argumento no usado en la demostración de la primera, esa observación no sería permisible.

En la demostración del siguiente teorema, que es también una equivalencia, en vez de desentendernos de la segunda implicación mediante una observación, bosquejamos la demostración sin justificar el desarrollo del argumento.

EDITORA AMERICA UNIDA
Unión Central 1043 - Fono 721374
SANTIAGO

INTRODUCCION A LA LOGICA SIMBOLICA

Teorema 10. $0 < -x$, *si y sólo si* $x < 0$.

Demostración. Si $x < 0$, entonces, por el Axioma 13,

$$-x + x < -x + 0,$$

y de aquí que, aplicando el Axioma 6 al segundo miembro y el Teorema 7 y la ley conmutativa al primero, inferimos:

$$0 < -x.$$

De modo semejante, si $0 < -x$, entonces

$$x + 0 < x + (-x)$$

y así

$$x < 0. \qquad \text{Q.E.D.}$$

Con respecto del enunciado de los Teoremas 9 y 10, hay otro estilo con el cual debe familiarizarse el lector. El Teorema 9 podría redactarse como sigue:

A fin de que $-y < -x$, *es necesario y suficiente que* $x < y$.

La frase 'necesario y suficiente' desempeña el mismo papel que 'si y sólo si' desempeña en el enunciado original. 'Si $x < y$' es la condición suficiente y 'sólo si $x < y$' es la condición necesaria. (Este giro idiomático se discutió en el § 1.3).

El siguiente teorema justifica la operación de sustracción.

Teorema 11. *Existe exactamente un número z tal que* $x = y + z$.

Demostración. Primeramente demostramos que existe cuando menos una z que satisface el teorema. En virtud del Teorema 7

(1) $$y + (-y) = 0,$$

de donde

$$(y + (-y)) + x = 0 + x,$$

y así, por la ley conmutativa y el Axioma 6

(2) $$x = (y + (-y)) + x.$$

Aplicando a (2) la ley asociativa, se obtiene:

$$x = y + ((-y) + x),$$

y claramente $(-y) + x$ es una z satisfactoria.

La demostración de que existe cuando menos una z es exacta-

mente igual a la parte similar de la demostración del Teorema 5; considérese una z' tal que $x = y + z'$; entonces

$$y + z = y + z',$$

y así, por la ley de la cancelación y la adición

$$z = z'. \qquad \text{Q.E.D.}$$

En esta demostración se sigue claramente de (2) que $(-y) + x$ es un número que satisface la condición deseada; esto es, si $z = (-y) + x$, entonces $x = y + z$. En estas diversas demostraciones que sólo entrañan adición y no multiplicación, nos remitimos a las leyes conmutativas y asociativas; se entiende que nos referimos a las leyes para la operación de adición. La estrategia de esta demostración consiste en escoger el axioma más cercano al teorema (Axioma 8) y trabajar a partir del mismo. No había necesidad de usar la operación negativa en esta prueba; es suficiente tener en (1) que existen números u y v con $x + u = 0$ y $y + v = 0$. Se introdujo la operación solamente para dar más práctica en su manejo.

DEFINICION 2. *$x - y = z$, si y sólo si $x = y + z$.*

Ha sido seguida la notación tradicional aritmética al usar el mismo signo '—' en las Definiciones 1 y 2, aunque en la Definición 1 denota una operación unaria (esto es, una función de un argumento) y en la Definición 2 denota una operación binaria (esto es, una función de dos argumentos). Esta confusión en la notación tipo está más o menos justificada por el teorema siguiente:

TEOREMA 12. *$x - y = x + (-y)$.*

DEMOSTRACIÓN. En virtud del Teorema 11 hay una z única tal que

(1) $$x = y + z$$

y de aquí, por la definición 2

$$x - y = z.$$

Las siguientes identidades demuestran que $x + (-y)$ es igual a z y de aquí que sea igual a $x - y$:

$$\begin{aligned}
x + (-y) &= (-y) + x && \text{(Ley conmutativa)}\\
&= (-y) + (y + z) && \text{(Por (1))}\\
&= ((-y) + y) + z && \text{(Ley asociativa)}\\
&= (y + (-y)) + z && \text{(Ley conmutativa)}\\
&= 0 + z && \text{(Teorema 7)}\\
&= z + 0 && \text{(Ley conmutativa)}\\
&= z && \text{(Axioma 6)}
\end{aligned}$$

Q.E.D.

Los tres teoremas siguientes ejemplifican algunas típicas y útiles propiedades de la relación $<$.

TEOREMA 13. *No es verdad que $x < x$.*

DEMOSTRACIÓN. El teorema se sigue inmediatamente del Axioma 10, haciendo $x = y$. Q.E.D. Esta demostración no es tan obvia que pueda ser despachada en una oración Si damos a 'y' el valor 'x' en el Axioma 10, obtenemos la fórmula

Si $x < x$, entonces no es verdad que $x < x$,

lo que lógicamente es equivalente a:

No es verdad que $x < x$.

Podría también darse una breve demostración indirecta suponiendo que existe una x tal que $x < x$ y a continuación aplicar el Axioma 10 para deducir una contradicción. Adviértase el uso de la expresión 'haciendo $x = y$', para indicar una aplicación de *US*.

TEOREMA 14. *Es válida exactamente una de las aserciones siguientes: $x = y, x < y, y < x$.*

DEMOSTRACIÓN. Se sigue claramente del Axioma 12 que cuando menos una de las tres fórmulas es verdadera. Demostramos que a lo sumo una es verdadera, demostrando que dos no son verdaderas simultáneamente. Primero, si $x = y$ y $x < y$, se obtiene inmediatamente: $x < x$, lo que contradice el Teorema 13 y es obvio que exactamente el mismo argumento demuestra que no puede ser que simultáneamente

$$x = y \quad \text{y} \quad y < x.$$

Finalmente, si fueran simultáneamente verdaderos $x < y$ y $y < x$, se contradiría el Axioma 10. Q.E.D.

La demostración del Teorema 14 ha sido nuestra primera ocasión de usar informalmente una demostración indirecta. Las dos aplicaciones en la demostración son demasiado obvias para requerir comentario.

TEOREMA 15. *Si $x < 0$ y $y < z$, entonces $x{\cdot}z < x{\cdot}y$.*

DEMOSTRACIÓN. Según la hipótesis del teorema y el Teorema 10,

$$0 < -x \quad \text{y} \quad y < z.$$

De aquí que por el Axioma 14

$$(-x)\cdot y < (-x)\cdot z,$$

y en seguida por el Teorema 8

$$-(x \cdot y) < -(x \cdot z),$$

de donde, por el Teorema 9

$$x \cdot z < x \cdot y \qquad \text{Q.E.D.}$$

Esta prueba ejemplifica el método tipo cuando el teorema que ha de ser probado es una oración condicional: supóngase la hipótesis y véase si la conclusión puede obtenerse por métodos directos.

El teorema siguiente es la ley de la cancelación para la multiplicación. La demostración de esta ley es la más intrincada dada hasta ahora y debe estudiarse con cierto cuidado. Nótese que es una prueba indirecta; el método directo tipo que se acaba de mencionar no es aplicable a este teorema.

TEOREMA 16. *Si* $x \neq 0$ *y* $x{\cdot}y = x{\cdot}z$, *entonces* $y = z$.

DEMOSTRACIÓN. Damos una prueba indirecta. Supóngase $y \neq z$. Entonces, por el Axioma 12, $y < z$ o $z < y$. Puesto que la hipótesis del teorema es simétrica por lo que respecta a y y a z, bastará con demostrar el teorema en el supuesto de que

$$(1) \qquad y < z.$$

Por la hipótesis del teorema, $x \neq 0$; de aquí que, por el Axioma 12, $0 < x$ o $x < 0$, lo que da dos casos que deben considerarse separadamente. Si $0 < x$, entonces, según (1) y el Axioma 14, se tiene:

$$x \cdot y < x \cdot z,$$

que juntamente con el Teorema 14 contradice la hipótesis de que $x{\cdot}y = x{\cdot}z$. Por otra parte, si $x < 0$, entonces en virtud de (1) y del Teorema 15

$$x \cdot z < x \cdot y,$$

que juntamente con el Teorema 14 contradice la hipótesis de que $x{\cdot}y = x{\cdot}z$ Q.E.D.

En esta demostración, en diversas etapas de la misma, tenemos disyuntivas que considerar. Los matemáticos no ven con agrado demostraciones que requieran subdivisión en un gran número de casos, pero la división en dos o tres casos es frecuente y a veces inevitable. Al comienzo se presenta la disyuntiva: $y < z$ o $z < y$. Pero no necesitamos estudiar estos dos casos separadamente, pues el papel de x y el de z son exactamente iguales en la hipótesis. Para

la demostración del otro caso, puede ser repetida la demostración del primer caso, intercambiando las letras. Cuando nos enfrentamos a las alternativas: $0 < x$ o $x < 0$, la situación es diferente. La demostración de un caso $(0 < x)$ depende del Axioma 14 y la demostración del otro caso $(x < 0)$, depende del Teorema 15. Cuando en una etapa de una demostración surge una variedad de posibilidades, la demostración no puede progresar hasta que ha sido investigada en su génesis y satisfactoriamente eliminada cada una de estas posibilidades (esto es, disyuntivas). Por supuesto, en muchos casos el número de las posibilidades puede ser reducido en forma inmediata por consideraciones de simetría, como al comienzo de esta demostración. La cuestión medular que debe recordarse es que si se suscitan casos disyuntivos, debe indicarse en la demostración algún método de eliminarlos.

En este punto concluimos el análisis de las pruebas informales. El desarrollo sistemático de la aritmética se lleva un tanto más adelante en los ejercicios que acompañan este apartado.

EJERCICIOS

1. Formúlese el Teorema 5 en notación lógica. (Este ejercicio lleva el propósito de aclarar cualquier remanente de confusión sobre el exacto significado de la expresión 'existe exactamente un...').

2. Demuéstrense las siguientes aserciones, relativas a la operación negativa y a la operación de sustracción.

(a) $-x = (-1)\cdot x$
(b) $(-x)\cdot y = x\cdot(-y)$
(c) $(x - y) + (y - z) = x - z$
(d) $x\cdot(y - z) = x\cdot y - x\cdot z$
(e) $(x - y) + (z - w) = (x + z) - (y + w)$
(f) $(x - y) - (z - w) = (x + w) - (y + z)$
(g) $(x - y)\cdot(z - w) = (x\cdot z + y\cdot w) - (x\cdot w + y\cdot z)$

3. Aplicando la definición

$$x \leq y, \quad \textit{si y sólo si,} \quad x = y \quad \textit{o} \quad x < y,$$

demuéstrese que

$$\textit{si} \quad x + y \leq x + z, \quad \textit{entonces} \quad y \leq z.$$

4. Demuéstrese lo siguiente:

(a) Si $x + x = 0$, entonces $x = 0$.
(b) Si $x \cdot x = 0$, entonces $x = 0$.
(c) Si $u \neq 0$ y $v \neq 0$, entonces $u \cdot v \neq 0$.

5. Demuéstrese lo siguiente:

(a) Si $0 < x$ y $x \cdot y < x \cdot z$, entonces $y < z$.
(b) Si $x < 0$ y $x \cdot y < x \cdot z$, entonces $z < y$.
(c) Si $0 < x$, $y < 0$, y $z < 0$, entonces $x \cdot y < y \cdot z$.
(d) Si $x < 0$, $0 < y$, y $0 < z$, entonces $x \cdot y < y \cdot z$.
(e) $x - y < x - z$, si y sólo si $z < y$.

6. Defina la operación de tomar el valor absoluto de un número y demuestre que:

(a) $|x \cdot y| = |x| \cdot |y|$
(b) $|x + y| \leq |x| + |y|$
(c) $|x| - |y| \leq |x - y|$
(d) $x \cdot |y| \leq |x \cdot y|$
(e) $|x - y| \leq |x - z| + |y - z|$

CAPITULO 8

TEORIA DE LA DEFINICION

§8.1. Ideas tradicionales. En la lógica antigua, una definición es la delimitación de una especie, mediante el enunciado del género que la incluye y la diferencia específica, o característica distintiva de la especie. Un ejemplo típico es la definición de hombre como animal racional. El género es el género animal y la característica distintiva es la racionalidad. (Lo que ha sido enunciado en forma compacta, es la teoría aristotélica de la definición). Muchos textos que no son tan tradicionales como para exigir una adhesión estricta al análisis aristotélico, seriamente promulgan las cuatro "reglas" tradicionales de la definición:

1. Una definición debe dar la esencia de lo definido.
2. Una definición no debe caer en un círculo vicioso.
3. Una definición no debe ser negativa cuando puede ser positiva.
4. Una definición no debe ser expresada en lenguaje figurado u oscuro.

Ciertamente, estas reglas tienen una aplicación efectiva como normas prácticas. Prohíben usar como definiciones enunciados como el siguiente:

> La belleza es la eternidad mirándose en un espejo—
> KHALIL GIBRAN, *El Profeta*,

que viola la Regla 4, o

> Fuerza no es un concepto cinemático,

que viola la Regla 3. Pero por otra parte estas reglas son de poca ayuda para aclarar el concepto formal de una definición apropiada, dentro de una teoría enunciada con precisión, como la teoría de los números reales, parcialmente desarrollada en el último capítulo.

Por ejemplo, podemos definir en aritmética la pseudooperación $*$ como sigue:

(1) $x * y = z$, si y sólo si $x < z$ y $y < z$.

Aplicando nuestro conocimiento intuitivo de la aritmética, podemos fácilmente usar esta pseudooperación para deducir una contradicción

$$1 * 2 = 3$$

puesto que

$$1 < 3 \text{ y } 2 < 3;$$

pero también

$$1 * 2 = 4,$$

puesto que

$$1 < 4 \text{ y } 2 < 4.$$

De aquí inferimos

$$3 = 4,$$

que contradice el hecho familiar de que $3 \neq 4$. Pero es el caso que todas las cuatro reglas tradicionales parecen satisfechas por (1). Cualquier clase de entidades que sean las esencias, ciertamente la esencia de la operación $*$ es enunciada por (1). Además, la definición no cae en un círculo vicioso, no es negativa, oscura, ni está en lenguaje figurado. Resulta transparentemente claro de este sencillo ejemplo que, a fin de desarrollar una adecuada teoría formal de la definición, debemos penetrar más allá del nivel de evidencia un tanto trivial de las cuatro reglas tradicionales.

§8.2. Criterios para definiciones adecuadas. Una definición tradicional *per genus et differentiam* es llamada a menudo una definición *real*, porque se dice que caracteriza la esencia de una especie. La clase de definición común en matemáticas, esto es, definiciones que introducen un nuevo símbolo, son a menudo llamadas definiciones *verbales* o *nominales*. Sin embargo, no está claro cómo puede hacerse una distinción tajante entre ambas clases de definiciones. Para nuestro propósito, basta con comprender que una definición es un enunciado que establece el significado de una expresión. La definición lleva esto a efecto relacionando la expresión que define (el *definiendum*) con otras expresiones (el *definiens*) de que ya se dispone.

Cuando menos dos cuestiones se suscitan inmediatamente de este vago enunciado sobre lo que son las definiciones. ¿Qué significa 'otras expresiones de que ya se dispone'? ¿Qué restricciones, si las hay, se imponen a la forma lógica de las oraciones que pueden servir como definiciones? La respuesta a la primera pregunta es que pensamos en la introducción de una definición dentro de una teoría

específica, como la teoría elemental de la aritmética considerada en el capítulo anterior. Como aquí se entiende, una *teoría* se caracteriza en función de sus símbolos y axiomas primitivos y no lógicos. En el caso de la teoría del capítulo anterior, los símbolos primitivos son el símbolo de relación '$<$', el símbolo de operación '+' y '·', así como las constantes '0' y '1'. Los axiomas son justamente los quince axiomas dados al principio de este capítulo. La teoría de grupos introducida en el Cap. 5 tiene tres símbolos primitivos y tres axiomas. En el presente capítulo, se supone una lógica de predicados de primer orden, con identidad, tal como se ha desarrollado en los capítulos anteriores, y sólo estudiamos teorías que puedan formalizarse dentro del marco de esta lógica.

La primera definición de una teoría es, en consecuencia, una oración de cierta forma, que establece el significado de un nuevo símbolo de la teoría en función de los símbolos primitivos de la misma. La segunda definición de una teoría es una oración de cierta forma que establece el significado de un segundo símbolo de la teoría, en función de los símbolos primitivos y del primer símbolo definido. Y de modo semejante para las definiciones subsecuentes. El particular que debe advertirse es que las definiciones de una teoría se introducen una por una, en cierta secuencia determinada. A causa de esta secuencia determinada, podemos siempre referirnos *significativamente* a las definiciones *precedentes* de la teoría. A menudo es conveniente adoptar el punto de vista de que cualquier símbolo definido debe definirse en función solamente de los símbolos primitivos de la teoría. En este caso no hay necesidad de introducir definiciones en una secuencia determinada. Sin embargo, la práctica común en matemáticas es usar símbolos previamente definidos para definir los nuevos símbolos y para dar una razón exacta de esta práctica, se necesita una secuencia determinada de definiciones.

Se hizo la observación en el capítulo anterior de que desde el punto de vista de la lógica de la inferencia una definición en una teoría se considera simplemente como un nuevo axioma o premisa. Pero no se quiere decir que una definición reforzará la teoría en forma sustancial. El motivo de introducir un nuevo símbolo es el de facilitar la investigación deductiva de la estructura de la teoría, pero no agregar algo a esa estructura. Dos criterios que hacen más específicas estas ideas intuitivas sobre el carácter de las definiciones son que (i) un símbolo definido debe ser siempre eliminable de cualquier fórmula de la teoría, y (ii), una nueva definición no permite demostrar la existencia de relaciones entre los antiguos símbolos que hayan sido hasta el momento imposibles de demostrar; esto es, no

funciona como un axioma creador.[1] En el capítulo anterior, por ejemplo, introducimos el símbolo de sustracción para la equivalencia.

(1) $x - y = z$ si y sólo si $x = y + z$.

Podemos usar (1) para eliminar cualquier incidencia del símbolo de sustracción. Así, en virtud de (1) eliminamos '−' de:

Si $y \neq 0$, entonces $x - y \neq x$,

y obtenemos el enunciado aritméticamente equivalente

Si $y \neq 0$, entonces $x \neq y + x$.

Parece razonable exigir que cualquier definición que introduzca un nuevo símbolo pueda ser usada para eliminar todas las subsecuentes incidencias significativas de los nuevos símbolos. Ser eliminable es una propiedad característica de un símbolo definido, esto es, opuesto a un símbolo primitivo. Podemos ahora formalizar el concepto de eliminabilidad.

CRITERIO DE ELIMINABILIDAD. *Una fórmula* **S** *que introduce un nuevo símbolo de una teoría satisface el criterio de eliminabilidad si y sólo si cada vez que* $\mathbf{S}_1$ *es una fórmula en que el nuevo símbolo ocurre, entonces hay una fórmula* $\mathbf{S}_2$ *en que el nuevo símbolo no ocurre, de manera que* $\mathbf{S} \rightarrow (\mathbf{S}_1 \leftrightarrow \mathbf{S}_2)$ *es derivable de los axiomas y definiciones precedentes de la teoría.*

Como lo sugiere la redacción de este criterio, de aquí en adelante no llamamos nuevos axiomas a las definiciones, aunque funcionen como axiomas adicionales en la inferencia lógica. La razón de esta restricción terminológica es obvia. Al diseñar una nueva teoría para la investigación, deseamos enunciar al principio los axiomas creadores. Puesto que las definiciones son teóricamente dispensables, no deseamos darles la misma categoría que a los axiomas básicos de la teoría.

En el siguiente enunciado se formaliza el concepto de que una definición no es creadora.

CRITERIO DE NO CREATIVIDAD. *Una fórmula* **S** *que introduce un nuevo símbolo de una teoría satisface el criterio de no creatividad si y sólo si no hay una fórmula* **T** *en que el símbolo no aparezca, tal que* $\mathbf{S} \rightarrow \mathbf{T}$ *sea derivable de los axiomas y definiciones precedentes de la teoría, pero* **T** *no sea derivable.*

En otras palabras, no podemos permitir que una fórmula **S** que introduce un nuevo símbolo haga posible la deducción de un teorema,

[1] Estos dos criterios fueron formulados por primera vez por el lógico polaco S. Lesniewski (1886–1939); también él fue la primera persona que dio reglas de definición que satisfacían los criterios.

anteriormente improbable, enunciado totalmente en función de símbolos primitivos y de símbolos previamente definidos. Un ejemplo de fórmula que no satisface este criterio de no creatividad, es el segundo axioma para grupos, si consideramos una teoría más limitada que la de los grupos. El único símbolo primitivo de nuestra teoría es el símbolo binario '$\circ$', y el axioma único, el axioma asociativo:

(1) $$x \circ (y \circ z) = (x \circ y) \circ z.$$

Como primera definición de esta teoría proponemos ahora la siguiente fórmula, que introduce la nueva constante individual 'e':

(2) $$x \circ e = x.$$

Pero, aplicando el criterio de no creatividad, rechazamos (2) como definición propuesta en nuestra teoría, pues de (2) se puede derivar inmediatamente:

(3) $$(\exists y)(x)(x \circ y = x).$$

Notamos que (3) es una fórmula cuyo único símbolo no lógico es el símbolo primitivo de la teoría, pero es trivial encontrar una interpretación que demuestre que (3) no puede ser deducido de (1). Así pues, (2) es creadora y debe ser rechazada como definición adecuada.

Debe notarse que una consecuencia especial del criterio de no creatividad es el de consistencia relativa. Si los axiomas y las definiciones precedentes son consistentes y si una fórmula que introduce un nuevo símbolo puede usarse para deducir una contradicción, entonces la nueva fórmula no satisface el criterio de no creatividad, pues de una contradicción puede deducirse cualquier fórmula, puesto que $(\mathbf{P} \ \& \ -\mathbf{P}) \rightarrow \mathbf{Q}$ es una tautología. En consecuencia, no necesitamos como tercer criterio independiente el de consistencia relativa.

En la siguiente sección nos dedicamos a la tarea de establecer reglas de definición que garanticen la satisfacción de los dos criterios de eliminabilidad y no creatividad.

EJERCICIOS

1. Usese la definición de sustracción para eliminar el símbolo de sustracción de los siguientes enunciados.

(a) $x - 0 = x$.
(b) $x - 1 \neq 1$.
(c) Si $x \neq y$, entonces $x - y \neq y - x$.
(d) Si $0 < x$ y $0 < y$, entonces $x - y \neq x + y$.

2. Dada la teoría consistente en el axioma único sobre la operación binaria símbolo 'o':

$$x \circ y = y \circ x,$$

pruébese que la fórmula

$$x \circ y = e$$

que introduce la constante individual '*e*' creadora en esta teoría.

§8.3. Reglas de las definiciones adecuadas. En las teorías enunciadas en lenguaje preciso (sea la materia matemáticas puras, física o psicología), ordinariamente introducimos tres clases de símbolos definidos: símbolos de relación, símbolos de operación y constantes individuales. Así, en la teoría de la aritmética de los números reales comenzada en el capítulo anterior, '$\leq$' y '$\geq$' son símbolos de relación definidos, los símbolos de operación y sustracción son símbolos de operación definidos, y los nombres de cualesquiera símbolos, excepto 0 y 1 son constantes individuales definidas (por ejemplo, '2', '3' y '4').

Para mayor sencillez del enunciado, introduciremos reglas separadas para cada una de estas clases de símbolos. Primero enunciamos reglas que requieran que una definición adecuada sea una equivalencia. Subsecuentemente discutimos el uso de identidades para definir símbolos de operación y constantes individuales. Al tratar de definiciones que son equivalencias es costumbre introducir el nuevo símbolo en el lado izquierdo de la equivalencia y llamar a este lado el *definiendo*. El lado derecho se llama el *definiente*. Así, en la definición de sustracción '$x - y = z$' es el definiendo y '$x = y + z$' es el definiente. Usamos este lenguaje de 'definiendo' y 'definiente' en todo este capítulo.

> REGLA PARA DEFINIR SÍMBOLOS DE RELACIÓN. *Una equivalencia* **D** *que introduce un nuevo símbolo* **P** *de n posiciones, es una definición adecuada en una teoría si y sólo si* **D** *es de la forma* $\mathbf{P}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \leftrightarrow \mathbf{S}$ *y se satisfacen las siguientes restricciones:* (i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *son variables distintas;* (ii) **S** *no tiene variables libres distintas de* $\mathbf{v}_1, \ldots, \mathbf{v}_n$; *y* (iii) **S** *es una fórmula en que las únicas constantes no lógicas son símbolos primitivos y símbolos previamente definidos de la teoría.*

Nótese que el definiendo $\mathbf{P}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ es una fórmula atómica, forma necesaria para garantizar la eliminación del símbolo definido de relación de todo contexto posible. Correspondiendo a la convención establecida en el capítulo anterior, las variables $\mathbf{v}_1, \ldots, \mathbf{v}_n$ están libres en la equivalencia **D**. Podría obtenerse una estricta conformidad a las reglas formales de inferencia, poniendo a la cabeza cuantificadores universales. Algunos ejemplos acompañados de dis-

cusión, ayudarán a aclarar las tres restricciones de la regla. La exigencia de que las variables $v_1, \ldots, v_n$ sean distintas, impide definiciones como:

(1) $x \leq x$, si y sólo si $x = x$ o $x < x$.

La fórmula (1) no define realmente el símbolo binario de relación '$\leq$', puesto que solamente una variable ocurre en el definiendo. Con (1) a mano, no sabríamos cómo eliminar '$\leq$', de la fórmula $x \leq y$. El definiente de (1) debe ser considerado como que define un símbolo de relación unaria, digamos 'U' (las relaciones unarias son únicamente propiedades, esto es, una propiedad es una relación de una posición):

$U(x)$ si y sólo si $x = x$ o $x < x$.

Por supuesto, la propiedad U es una propiedad trivial universal, poseída por todo número. Como segundo ejemplo, considérese la definición de la relación cuaternaria que existe entre cuatro miembros, si la diferencia entre el primero y el segundo es menor que la existente entre el tercero y el cuarto. Usamos la letra 'A' como símbolo de relación.

(2) $A(x, y, u, v)$, si y sólo si $x - y < u - v$

La generalidad de (2) y, en consecuencia, la eliminabilidad del símbolo de relación 'A' se perdería si (2) fuera reemplazada por

(3) $A(x, y, u, x)$, si y sólo si $x - y < u - x$.

El definiente de (3) realmente define la relación ternaria T:

(4) $T(x, y, z)$, si y sólo si $x - y < z - x$.

[Nótese el significado intuitivo de la relación T:$T(x, y, z)$, justamente cuando x es menor que el medio de y y z (esto es, $x < (y + z)/2$).]

La segunda restricción impide definiciones como

(5) $R(x)$, si y sólo si $x + y = 0$.

Cuando se agrega (5) a los axiomas de aritmética, es posible deducir una contradicción. La fuente de la dificultad es la aparición de la variable 'y' en el definiente, pero no en el definiendo. Ahora bien, (5) es lógicamente equivalente a los dos enunciados:

(6) Si $x + y = 0$, entonces $R(x)$

(7) Si $R(x)$, entonces $x + y = 0$.

Pero de la lógica de cuantificadores sabemos que (6) es equivalente a

(8) Si hay una y tal que $x + y = 0$, entonces $R(x)$,

y (7) es equivalente a

(9) Si $R(x)$, entonces para toda y, $x + y = 0$.

De (8) y (10), inmediatamente inferimos la patente falsedad:

(10) Si hay una y tal que $x + y = 0$, entonces para toda y, $x + y = 0$.

[Nótese que la variable 'x' se deja libre en esta discusión, pues aparece de una manera apropiada en el definiendo y en el definiente de (5)].

Por otra parte la segunda restricción no impide que las variables estén libres en el definiendo, pero no en la definición. En consecuencia, admitimos como definición adecuada:

(11) $Q(x, y)$, si y sólo si $x > 0$.

No hay razón formal definida para prohibir que las variables estén libres en el definiendo y no lo estén en el definiente. Por otra parte, esas variables no se usan para expresar algo intuitivamente significativo; la trivialidad del papel que desempeñan queda subrayada por el hecho de que siempre es posible encontrar una equivalencia que tenga el mismo contenido lógico y las mismas variables libres en el definiendo y el definiente. Así, podemos convertir (11) en una fórmula que tenga las mismas variables libres en el definiendo y en el definiente conjuntando al definiente la verdad lógica $y = y$. La nueva fórmula:

(12) $Q(x, y)$, si y sólo si $x > 0$ & $y = y$

es lógicamente equivalente a (11). Puede usarse una conjunción similar de identidades lógicas para convertir cualesquiera equivalencias que tengan mas variables libres en el definiendo que en el definiente, en una que tenga el mismo número en ambos.

La tercera restricción simplemente prohíbe dos clases de circularidad de definición. No se puede admitir como definición adecuada:

(13) $R(x)$, si y sólo si $R(x)$;

una verdad lógica como la (13) no sería creadora. Su defecto estriba en que no satisface el criterio de eliminabilidad. La fórmula (13) no suministra un procedimiento para eliminar el símbolo de relación 'R'. De naturaleza similar son el par de equivalencias:

(14) $R(x)$, si y sólo si no es verdad que $P(x)$,

(15) $P(x)$, si y sólo si no es verdad que $R(x)$.

Si definimos el símbolo de relación 'R' en función del nuevo símbolo de relación 'P', y viceversa, entonces no podemos eliminar ninguno de los dos, en favor de la notación primitiva. Así pues, se tiene la exigencia de que no aparezca en la definición ningún símbolo nuevo y de que las definiciones se den en una secuencia fija.

Pasamos ahora a la regla para definir símbolos. Tiene que añadirse una restricción esencialmente nueva a las tres que son necesarias para los símbolos de relación. Al enunciar la regla, usamos la notación tipo (standard)

$$(\mathrm{E!}\mathbf{w})\mathbf{S}$$

para simbolizar

Hay exactamente una **w** *tal que* **S**

REGLA PARA DEFINIR SÍMBOLOS DE OPERACIÓN. *Una equivalencia* **D** *que introduce un nuevo símbolo de operación de nueve posiciones* $\mathbf{O}$, *es una definición adecuada en una teoría si y sólo si* **D** *es de la forma*

$$\mathbf{O}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \mathbf{w} \leftrightarrow \mathbf{S},$$

y se satisfacen las siguientes restricciones: (i) $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}$ *son variables distintas*, (ii) **S** *no tiene más variables libres que* $\mathbf{v}_1 \ldots, \mathbf{v}_n, \mathbf{w}$, (iii) **S** *es una fórmula en que las únicas constantes no lógicas son símbolos primitivos y símbolos previamente definidos de la teoría, y* (iv) *la fórmula* $(\mathrm{E!}\mathbf{w})\mathbf{S}$ *es derivable de los axiomas y símbolos precedentes de la teoría.*

La consideración de la pseudooperación $\star$ introducida en § 8.1 es suficiente para justificar la cuarta restricción. Mediante el uso de la pseudooperación $\star$ se pudo deducir una contradicción en § 8.1, justamente porque '$x \star y$' no designa una entidad única. En el caso de la definición de la operación $\star$:

(16) $x \star y = z$, si y sólo si $x < z$ y $y < z$

no podemos probar que hay exactamente una z tal que $x < z$ y $y < z$, y, en consecuencia, (16) no es una definición adecuada. La función de la cuarta restricción consiste en exigir que la definición de cualquier nueva operación sea precedida por un teorema que garantice que la operación está definida unívocamente. En el Cap. 7 las definiciones de la operación negativa y la sustracción estuvie-

ron ambas precedidas por esos teoremas justificativos. La generalidad de la regla definiente, que se aplica a operaciones de cualquier complejidad, no debe oscurecer el hecho de que generalmente nos ocupamos de operaciones binarias, cuyas definiciones son de la forma

$$x \circ y = z, \text{ si y sólo si } \mathbf{S}(x, y, z),$$

y necesitamos un teorema precedente en el sentido de que para toda x y toda y, hay exactamente una z tal que $S(x, y, z)$.

Podemos considerar a las constantes individuales como símbolos de operación de rango cero. Sin embargo, a causa de la forma un tanto desconcertante de las definiciones de constantes individuales cuando estas definiciones son equivalencias más bien que identidades, parece aconsejable enunciar explícitamente la regla. La razón para insistir en la consideración de definiciones de constantes individuales en forma de equivalencias es que no siempre es posible introducirlas por medio de identidades, como veremos dentro de poco. A fin de hacer intuitivamente clara la regla original, podemos primero indicar cómo definiríamos las constantes '0' y '1' si hubiéramos formulado nuestros axiomas sin usarlos como símbolos primitivos. Introduciríamos las dos equivalencias siguientes:

(17) $0 = y$, si y sólo si para toda x, $x + y = x$,

(18) $1 = y$, si y sólo si para toda x, $x \cdot y = x$.

Nótese que la variable 'x' está ligada en el definiente de (17) y en el de (18). La restricción sobre el carácter único de y debe aplicarse a las constantes individuales. Si se omitiera la restricción, podríamos introducir una constante 'b' por medio de una definición como la siguiente:

(19) $b = y$, si y sólo si $y > 0$

y deducir una contradicción, pues de (19) se sigue que

$$b = 1$$

y

$$b = 2.$$

De aquí que,

$$1 = 2,$$

que es absurdo.

Regla para definir constantes individuales. *Una equivalencia* **D** *que introduce una nueva constante individual* **c**, *es una definición adecuada en una teoría, si y sólo si* **D** *es de la forma*

$$\mathbf{c} = \mathbf{w} \leftrightarrow \mathbf{S},$$

y se satisfacen las siguientes restricciones: (i) **S** *no tiene variable libre distinta de* **w**, (ii) **S** *es una fórmula en que las únicas constantes no lógicas son símbolos primitivos y símbolos previamente definidos de la teoría, y* (iii) *la fórmula* (E!**w**)**S** *es deducible de los axiomas y definiciones precedentes de la teoría.*

En la sección siguiente se introducen reglas para definir símbolos de operación y constantes individuales por medio de identidades en vez de equivalencias. En los ejercicios se encontrarán varias aplicaciones de las reglas enunciadas en este apartado.

EJERCICIOS.

1. ¿Cuáles de las siguientes definiciones de símbolos de relación son impropias? En caso de ser impropias, diga qué restricción se viola.

(a) $R(x, y)$ si y sólo si hay una z tal que $x + y = z$.
(b) $R(x, y)$, si y sólo si $x + y > z$.
(c) $R(x, y)$, si y sólo si $y > 2$.
(d) $R(x, y)$, y $x \neq y$, si y sólo si $x + y > 0$.
(e) $R(x, x)$, si y sólo si $x + 1 > 0$.

2. ¿Cuáles de las siguientes definiciones de símbolos de operación son impropias? Si una definición es impropia, úsela para deducir intuitivamente una contradicción.

(a) $x \circ y = z$, si y sólo si $x < 2$ y $y < 3$ y $z < 4$.
(b) $x \circ y = z$, si y sólo si $x = 1$ y $y = 2$ y $z = 7$.
(c) $x \circ y = z$, si y sólo si $x < 1$ y $y < 2$ y $z = 7$.
(d) $x \circ y = z$, si y sólo si $x = 1$ y $y = 2$ y $z < 7$.
(e) $x \circ y = z$, si y sólo si $x + y = x + z$.
(f) $x \circ y = z$, si y sólo si $x \cdot z = y$.

3. ¿Cuál de las siguientes definiciones de constantes individuales es inadecuada? Si es inadecuada, indique qué restricción se viola.

(a) $c = y$, si y sólo si hay una x tal que $x + y = 1$.
(b) $c = y$, si y sólo si para toda x, si $x > 0$, entonces $x + y > 0$.
(c) $c = y$, si y sólo si $y = 1 + 1$.
(d) $c = y$, si y sólo si $y = y$.
(e) $c = y$, si y sólo si $y \neq y$.

4. Dé un ejemplo de definición inadecuada cuyo definiente es:

(a) Una verdad lógica.
(b) La negación de una verdad lógica.
(c) Un axioma de la teoría en que se propone la definición (constrúyase una teoría sencilla para este ejercicio).

5. Dése un ejemplo de definición adecuada, cuyo definiente sea:

(a) Una verdad lógica.
(b) La negación de una verdad lógica.
(c) La negación de un axioma de la teoría en que ocurre la definición (constrúyase una teoría sencilla para este ejercicio).

§8.4. Definiciones que son identidades. En el apartado anterior hicimos notar que a menudo son introducidos símbolos de operación y constantes individuales por medio de identidades en vez de por equivalencias. Deseamos ahora establecer reglas adecuadas para esas identidades.

Puesto que la definición

(1) $$2 = 1 + 1$$

parece más natural y elegante que la definición

(2) $$2 = y, \text{ si y sólo si } y = 1 + 1,$$

es razonable preguntar por qué se usan siempre equivalencias para definir las constantes individuales. La respuesta es muy sencilla: las solas identidades no son adecuadas para esa aplicación. Si se eliminaran '0' y '1' como símbolos primitivos de aritmética, podrían introducirse por medio de las equivalencias (17) y (18) del apartado anterior, pero estas equivalencias no podrían ser eliminadas en favor de identidades.[2]

Las mismas observaciones son aplicables al uso de equivalencias en vez de identidades para definir símbolos de operación. Por ejemplo, si no han sido dadas otras definiciones en la teoría de la aritmética del Cap. 7, entonces ni la operación negativa, ni el símbolo de operación de sustracción pueden ser definidos por una identidad; pero dado uno, puede ser definido el otro:

$$x - y = x + (-y),$$

$$-x = 0 - x.$$

O si la constante individual '-1' se define primero por medio de la equivalencia

$$-1 = x, \text{ si y sólo si } x + 1 = 0,$$

entonces el símbolo de operación negativa puede ser definido por la identidad:

$$-x = 0 - 1 \cdot x.$$

Como lo ilustran estos ejemplos, la posibilidad de definir un sím-

---

[2] Este enunciado sobre la inadecuación de identidades sería falso si nuestra lógica básica hubiera sido extendida hasta incluir un operador de descripción, 'el objeto x tal que...' Esta operación se simboliza usualmente por medio de '(ιx)', una notación introducida por Peano. Con este operador disponible, podríamos introducir una definición de '0':

$$0 = (\iota y)[(x)(x + y = x)].$$

El operador de descripción no fue introducido porque no es realmente necesario en la investigación deductiva de teorías formalizables dentro de la lógica de predicados de primer grado y porque su introducción complicaría más nuestras reglas básicas de inferencia.

bolo de operación por medio de una identidad, está en relación con el carácter exacto de las definiciones precedentes de la teoría.

Pasamos ahora a las reglas formales. En el caso de identidades que se usan como definiciones, llamamos al lado izquierdo de la identidad el definiendo, y al lado derecho el definiente, como podría esperarse de nuestro anterior procedimiento; aquí el definiendo y el definiente son ambos términos, en el sentido del Cap. 3. Correspondiendo a la noción de fórmula atómica introducida en el Cap. 3, es conveniente introducir la noción de *término atómico*. Un término atómico es un término que es una constante individual o que tiene exactamente una incidencia de un símbolo de operación. Así, '$x \cdot y$', '$x + y$', '$x - y$' y '1' son atómicos, mientras que '$x + (x + z)$' y '$x \cdot (y + z)$' no lo son. Cuando una definición es una identidad, se requiere que el definiendo sea un término atómico.

REGLA PARA DEFINIR SÍMBOLOS DE OPERACIÓN. *Uua identidad* **D** *que introduce un nuevo símbolo de operación* O *de n posiciones, es una definición adecuada en la teoría si y sólo si* **D** *es de la forma*

$$O(v_1, \ldots, v_n) = t,$$

y se satisfacen las siguientes restricciones: (i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *son variables distintas,* (ii) *el término* **t** *no tiene variables libres distintas de* $\mathbf{v}_1, \ldots, \mathbf{v}_n$, (iii), *las únicas constantes no lógicas del término* **t** *son símbolos primitivos y símbolos previamente definidos de la teoría.*

Es conveniente advertir que cuando se usan identidades para definir símbolos de operación no se necesitan teoremas justificativos para garantizar que el símbolo de operación está bien definido, pues la fórmula $(E_!\mathbf{w})$ $(\mathbf{t} = \mathbf{w})$ es una verdad de lógica.

Puesto que la regla para definir constantes individuales mediante el uso de una identidad es muy semejante a la que se acaba de dar para los símbolos de operación, su enunciado formal se deja como un ejercicio. Las definiciones tipo (standard) de nombres de números ejemplifican la introducción de constantes individuales por medio de identidades.

$$2 = 1 + 1$$
$$3 = 2 + 1$$
$$4 = 3 + 1$$
$$\cdots\cdots\cdots$$

EJERCICIOS

1. Sin suponer definiciones precedentes en la teoría de la aritmética del Cap. 7, ¿cuál de las siguientes operaciones aritméticas puede ser definida por medio de identidades? Cuando la respuesta sea positiva, enuncie la definición.

(a) Cuadrado de un número.
(b) Tercera potencia de un número.
(c) Valor absoluto de un número.
(d) Operación de añadir cinco a un número.
(e) Operación de multiplicar un número por dos.

2. Dense ejemplos que justifiquen las tres restricciones impuestas a la regla para definir símbolos de operación por medio de identidades.
3. Enuncie formalmente la regla para introducir constantes individuales por medio de definiciones que son identidades.
4. Dados como axiomas todos los enunciados verdaderos de aritmética que no contengan el símbolo '0', ¿es posible definir '0' por medio de una identidad?
5. Dados como axiomas todos los enunciados verdaderos de aritmética que incluyen los símbolos primitivos '+', '·', '<' y '0', ¿es posible definir '1' por medio de una identidad?
6. Dados como axiomas todos los enunciados verdaderos de aritmética que contienen los símbolos primitivos '+', '<', '0' y '1', ¿es posible definir por medio de una identidad el símbolo de multiplicación?

§8.5. El problema de la división entre cero. Que no todo esté orientado hacia lo mejor en éste, el mejor de los mundos posibles, ni siquiera en matemáticas, queda bien ilustrado por el acuciante problema de definir la operación de división en la teoría elemental de la aritmética. Si introducimos la definición.

(1) $x/y = z$, si y sólo si $x = y{\cdot}z$,

notamos inmediatamente que (1) no satisface la cuarta restricción impuesta a las equivalencias que definen símbolos de operación. Porque no podemos probar que dados dos números x y y hay una z única, tal que $x = y{\cdot}z$. Por ejemplo, no hay una z tal que $1 = 0{\cdot}z$, y cualquier número z tiene la propiedad de que $0 = 0{\cdot}z$.

Una modificación obvia de (1) es:

(2) $x/y = z$, si y sólo si $y \neq 0$ y $x = y{\cdot}z$

Pero no podemos probar que dadas dos x y y cualesquiera hay una z única que $y \neq 0$ y $x = y{\cdot}z$. Para ver esto, necesitamos simplemente considerar el caso de $y = 0$.

La aparente naturalidad de (2) sugiere un debilitamiento de la cuarta restricción, hasta dejarla en el requerimiento de que *cuando más* exista una z tal que $y \neq 0$ y $x = y{\cdot}z$. Es fácil demostrar que (2) satisface esta restricción debilitada, pero infortunadamente po-

demos derivar una contradicción de (2). La inferencia se desarrolla como sigue: Es una verdad lógica que

(3) $$\frac{1}{0}=\frac{1}{0}.$$

De aquí que exista una x tal que

(4) $$\frac{1}{0}=x,$$

y, en consecuencia, según (2)

$$0 \neq 0 \text{ y } 1 = 0 \cdot x,$$

lo que es absurdo. Debe tomarse nota de la situación de (3). Es un supuesto fundamental de la lógica desarrollada en este libro que todos los términos designan objetos. Dado un símbolo de operación, podemos siempre formular el análogo de (3) como una verdad lógica y a continuación inferir la clase de enunciado existencial que representa (4). Esas inferencias se discutieron en el Cap. 5 y no es necesario volver a examinar aquí su validez lógica. La burda posibilidad de cambiar la base lógica para resolver el problema de la división entre cero, juntamente con otras posibles soluciones, se menciona en el § 8.7.

A pesar de las dificultades que nos acosan, hay una manera formalmente satisfactoria de definir la división entre cero. Es una definición llamada libre de axiomas, pues no requiere un teorema previo para justificarla. Requiere que cuando $y = 0$ entonces $x/y = 0$. A causa de los diversos cuantificadores que ocurren en el definiente, usamos notación lógica para enunciar la definición.

(5) $$x/y = z \leftrightarrow (z')[z' = z \leftrightarrow x = y \cdot z']$$
$$\vee [-(\exists w)(z')(z' = w \leftrightarrow x = y \cdot z') \mathbin{\&} z = 0].$$

El carácter complicado de (5) arguye fuertemente a favor de otra solución, que tratamos en el apartado siguiente. Aunque (5) es bromoso, representa un acercamiento que teóricamente es importante al proporcionar un método para definiciones dadas de símbolos de operación, los cuales no dependen de teoremas previamente establecidos. La formulación general, de la cual (5) es un caso particular, es la siguiente.

$$\mathrm{O}(x_1, \ldots, x_n) = y \leftrightarrow (z)[z = y \leftrightarrow \mathrm{S}(x_1, \ldots, x_n, z)]$$
$$\vee [-(\exists w)(z)(z = w \leftrightarrow \mathrm{S}(x_1, \ldots, x_n, z) \mathbin{\&} y = 0].$$

Algunas veces la posición de $y = 0$, cuando la operación no está

"definida" en el sentido usual, parece un poco extraña. Así, muchos matemáticos se sentirían intranquilos viendo

$$\frac{2}{0} = 0,$$

y así, digamos:

$$1 + \frac{2}{0} = 1.$$

Estas cuestiones se discutirán ulteriormente en el § 8.7. Baste con decir aquí que parece que no hay un método de manejar la división entre cero que sea uniformemente satisfactorio.

EJERCICIOS

1. Dése una definición libre de axiomas de la operación de extracción de raíz cuadrada.
2. Pruébese que hay exactamente una z que satisface el definiente de (5).
3. Dése una definición libre de axiomas de la definición.

§8.6. Definiciones condicionales. La práctica usual en matemáticas es la de usar definiciones condicionales en vez de las burdas definiciones libres de axiomas presentadas al principio del apartado precedente. La técnica de la definición condicional consiste en hacer que una hipótesis preceda a una definición adecuada común. Así, una posible definición condicional de la división es:

(1) Si $y \neq 0$, entonces $x/y = z$, si y sólo si $x = y{\cdot}z$.

La principal desventaja de las definiciones condicionales estriba en que no satisfacen plenamente el criterio de eliminabilidad. Por ejemplo, si usamos (1) para definir la división, entonces no podemos eliminar el símbolo de división de la oración

$$\frac{1}{0} = \frac{1}{0}.$$

En cambio, es evidente que podemos eliminarlo en todos los casos "interesantes", esto es, todos los casos que satisfacen la hipótesis de (1).

Aunque las definiciones condicionales se usan tanto para definir símbolos de operación como de relación, enunciaremos únicamente las reglas sobre los símbolos de operación y dejaremos como ejercicio el caso de los símbolos de relación.

REGLAS PARA DEFINICIONES CONDICIONALES DE SÍMBOLOS DE OPERACIÓN. *Una implicación* **C** *que introduce un nuevo símbolo de operación* O *es una definición condicional en una teoría si y sólo si* **C** *es de la forma*

$$H \rightarrow [O(v_1, \ldots, v_n) = w \leftrightarrow S]$$

y se satisfacen las siguientes restricciones: (i) *la variable* **w** *no está libre en* **H**, (ii) *las variables* $\mathbf{v}_1, \ldots, \mathbf{v}_n$, **w** *son distintas,* (iii) **S** *no tiene variables distintas de* $\mathbf{v}_1, \ldots, \mathbf{v}_n$, **w** (iv) **S** y **H** *son fórmulas en que las únicas constantes no lógicas son símbolos primitivos y símbolos previamente definidos de la teoría,* y (v) *la fórmula* **H** → (E!**w**)**S** *es deducible de los axiomas y definiciones precedentes de la teoría.*

En vista de las varias explicaciones detalladas de reglas similares en secciones precedentes, no se necesitan comentarios específicos. En los ejercicios[3] se dan algunas explicaciones.

EJERCICIOS

1. Enuncie las reglas sobre definiciones condicionales de símbolos de relación.
2. Enuncie las reglas sobre las definiciones condicionales de constantes individuales.
3. Dé una definición condicional de la propiedad consistente en que un número sea impar. (SUGESTION: La hipótesis sería: x es un entero).
4. Dé una definición condicional de la propiedad consistente en que un número sea primo.
5. Dé una definición condicional de la operación de extraccion de raíz cuadrada.
6. Suponiendo que ya se han definido las operaciones exponenciales, dé una definición condicional de la operación logarítmica da base 10.

§8.7. Cinco métodos de la división por cero. Ya hemos mencionado varios aspectos del problema de la división por cero y hemos hecho la observación de que no existe una solución uniformemente satisfactoria. En este apartado queremos examinar cinco métodos de abordar el problema. El penúltimo de los cinco de la solución

[3] Un concepto filosófico relacionado con las definiciones condicionales, es el concepto de Rodolfo Carnap de una *oración de reducción* que suministra un método para relacionar predicados disposicionales, como "soluble en agua" a predicados directamente observables. Una posible oración de reducción para el predicado de solubilidad es:

Si x se coloca en agua, entonces x es soluble en agua si y sólo si x se disuelve.

Para ulteriores detalles véase el artículo de Carnap "Testability and Meaning", *Philosophy of Science*, Vol. 3 (1936) Págs. 419–471 y Vol. 4 (1937) Págs. 1–40.

que probablemente está más en consonancia con los procedimientos matemáticos ordinarios.

El primer método difiere de los demás en que recomienda un cambio en la lógica básica, para negar significado a expresiones cómo :

$$\frac{1}{0} = \frac{1}{0}.$$

Sin intentar caracterizar los cambios básicos, podemos, no obstante, oponer algunas objeciones básicas a este procedimiento. La primera objeción es que es indeseable complicar las reglas básicas de lógica, a menos que sea absolutamente necesario. En otras palabras, procede cambiar los fundamentos de la inferencia, sólo si otros procedimientos han fallado. En segundo lugar, si se adoptara ese cambio, la significación misma de las expresiones sería algunas veces difícil, si no es que imposible de dilucidar. Por ejemplo, supóngase que hemos añadido a nuestros axiomas del Cap. 7 axiomas suficientes para obtener los teoremas esperados sobre los números naturales (esto es, los enteros positivos). Considérese ahora la expresión:

(1) *Para todo número natural* $n, \frac{1}{n^*} = \frac{1}{n^*}$,

donde n^* es la operación unaria definida como sigue: $n^* = 1$. Si n es un entero impar, o si n es un entero par que es la suma de dos números primos; $n^* = 0$ si n es un entero par que no es la suma de dos números primos. El problema de la existencia de números enteros que no son la suma de dos primos es un famoso problema de matemáticas que todavía está sin solución (hipótesis de Goldbach). Así, sobre la base del primer procedimiento, la *significatividad* (no la verdad o falsedad) de (1) es una cuestión a dilucidar.

El segundo procedimiento es suponer que $x/0$ es un número real, pero definir la división mediante la definición condicional enunciada en el apartado anterior:

Si $y \neq 0$, entonces $x/y = z$, si y sólo si $x = y \cdot z$.

En este caso, para cada número, x, $x/0$ es un número real, pero no podemos probar qué número es. En realidad, no podemos ni siquiera decir sobre la verdad o falsedad del simple enunciado:

(2) $$\frac{1}{0} = \frac{2}{0}.$$

La incapacidad para probar la verdad o la falsedad de (2) es un

argumento contra el segundo procedimiento, pues deseamos que nuestros axiomas sean tan completos como sea posible.

El tercer procedimiento coincide con el segundo en hacer de $x/0$ un número real, pero difiere en que lo hace $x/0 = 0$ para todas las x. Esto elimina la indecisión de enunciados como (2). Para el tercer procedimiento, la definición apropiada de división es una definición adecuada similar a la definición libre de axiomas discutida al fin del § 8.5:

$$x/y = z \leftrightarrow [(y \neq 0 \rightarrow x = y \cdot z) \,\&\, (y = 0 \rightarrow z = 0)].$$

Una ventaja del tercer procedimiento es que permite la definición de cero por medio de una definición adecuada y directa que satisface plenamente los criterios de eliminabilidad y de no creatividad. La principal desventaja de este método es la mencionada en el § 8.5: muchos matemáticos se sienten desconcertados con la identidad:

$$\frac{x}{0} = 0.$$

El cuarto procedimiento concuerda con el segundo y tercero en que no requiere un cambio básico en la lógica; difiere en que coloca el objeto $x/0$ fuera del dominio de los números reales; más precisamente, difiere en que no hace posible probar que $x/0$ es un número real. La idea básica consiste en introducir un predicado 'R' que significa 'es un número real'. A todos los axiomas enunciados en el Cap. 7 deben añadirse hipótesis que usen ese predicado. Además, para garantizar que la adición y multiplicación de números dan números, debemos añadir los dos axiomas de oclusión

$$R(x) \,\&\, R(y) \rightarrow R(x + y),$$

$$R(x) \,\&\, R(y) \rightarrow R(x \cdot y).$$

La introducción del predicado 'R' ensancha nuestro dominio de individuos, pues ahora no se sigue que todo lo perteneciente al dominio sea un número real (en particular, que $1/0$ sea un número real). La introducción de 'R' tiene la consecuencia ulterior de que resulte natural hacer condicionales todas nuestras definiciones de relaciones y operaciones aritméticas. No tiene caso definir la relación de igual o menor que, por ejemplo, para cosas que no son números. De este modo, tendríamos:

$$[R(x) \,\&\, R(y)] \rightarrow [x \leq y \leftrightarrow (x = y \vee x < y)].$$

La definición de división sería:

(3) $$[R(x) \,\&\, R(y) \,\&\, R(z) \,\&\, y \neq 0] \rightarrow [x/y = z \leftrightarrow x = y \cdot z].$$

Si x es un número real, con (3) a la mano, podemos probar:

$$\frac{x}{0} \text{ es un número real}$$

y no podemos probar

$$\frac{x}{0} \text{ no es un número real}$$

pero ahora nos encontramos frente a la situación, contraria a la intuición, de llamar a $x/0$ un número real. La situación puede ser mejorada introduciendo en nuestro sistema un símbolo primitivo para algún objeto que no es un número real. Sin decir qué objeto es, designémoslo por 'ν'. Tenemos entonces el axioma

$$-R(\nu),$$

esto es, la aserción de que ν no es un número real. Podemos entonces definir la división por

(4) $[R(x) \& R(y)] \rightarrow [x/y = z$
$\leftrightarrow [(y \neq 0 \rightarrow (R(z) \& x = y \cdot z)) \& (y = 0 \rightarrow z = \nu)]].$

La virtud de (4) es que definitivamente coloca a $x/0$ fuera del dominio de los números reales para cualquier número x. Esa consecuencia parecería estar en completo acuerdo con el uso matemático ordinario. La definición (4) tiene también la virtud de hacer:

$$\frac{x}{0} = \frac{y}{0},$$

donde x y y son números reales, eliminando así una vasta proliferación de extrañas entidades matemáticas. Por otra parte, para ver de una vez por todas que no todo se orienta en el sentido de lo mejor en este mejor de los mundos posibles, nótese que si adoptamos (4), no podemos decidir si ν/ν es un número real. Si usamos definiciones condicionales e insistimos en no entremeternos con las leyes básicas de la identidad de los términos, entonces debemos estar preparados para la indecisoriedad del estado de entidades como ν/ν. Podemos tomar la actitud de que todo está en su debido orden en el dominio en que intentamos usar la operación de división y, en realidad, no nos importa lo que suceda en otra parte, si no pueden colarse inconsistencias.

Debe mencionarse un quinto procedimiento que es de considerable importancia, pero que no corresponde para nada a la práctica

matemática usual. La idea es sencilla: háganse desaparecer los símbolos de operación y las constantes individuales y úsense únicamente símbolos de relación. Así, '0' es reemplazado por el predicado primitivo 'Z' de una posición, donde se intenta que $Z(x)$ signifique que x es un elemento de identidad con respecto de la adición. El símbolo de operación ternaria '*A*' se usa para reemplazar el símbolo de la operación de adición:

$$A(x, y, z) \leftrightarrow x + y = z.$$

De modo semejante, el símbolo de operación ternaria '*M*' se usa para la multiplicación:

$$M(x, y, z) \leftrightarrow x \cdot y = z.$$

Con este aparato podemos fácilmente dar una definición adecuada del símbolo de relación de división:

$$D(x, y, z) \leftrightarrow -Z(y) \,\&\, M(y, z, x).$$

En este método no hay necesidad de entidades matemáticas desusadas, pero es extraordinariamente burdo trabajar continuamente con símbolos de relación, en vez de con símbolos de operación. Por ejemplo, la asociatividad de la adición tiene que ser expresada en alguna forma como la siguiente:

$$A(x, y, w) \,\&\, A(w, z, s_1) \,\&\, A(y, z, v) \,\&\, A(x, v, s_2) \rightarrow s_1 = s_2.$$

§8.8. **Principio de Padoa e independencia de los símbolos primitivos.** Cuando se dan los símbolos de una teoría, es natural preguntar si sería posible definir uno de ellos en función de los demás. El lógico italiano Alessandro Padoa formuló en 1900 un principio que aplica el método de interpretación y que puede ser usado para demostrar que los símbolos primitivos son independientes, esto es, que no puede uno ser definido en función de otro. El principio es sencillo: para demostrar que un símbolo primitivo dado es independiente a los restantes primitivos, encuéntrense dos interpretaciones de los axiomas de la teoría, tales que el primitivo dado tenga dos diferentes interpretaciones y los restantes símbolos primitivos tengan la misma interpretación. Por ejemplo, considérese la teoría de la preferencia basada en los primitivos símbolos de relación '*P*' (por preferencia estricta) e '*I*' (por indiferencia). Los axiomas de la teoría son:

A1. *si xPy & yPz, entonces* xPz.
A2. *si xIy & yIz, entonces* xIz.
A3. *Exactamente uno de los siguientes:*

$$xPy,\ yPx,\ xIy.$$

Deseamos demostrar que '*P*' es independiente de '*I*', esto es, no puede ser definido en función de '*I*'. Sea el dominio de interpretación para ambas interpretaciones el conjunto {1, 2}. Sea "*I*" interpretada como identidad en ambos casos. En un caso, sea '*P*' interpretado como '$<$' y en el otro como '$>$'. En la primera interpretación tenemos:

$$1 P 2$$

puesto que

$$1 < 2,$$

y consecuentemente, por el Axioma A3

$$\text{no} \quad 2 P 1.$$

Pero en la segunda interpretación tenemos:

$$2 P 1$$

puesto que

$$2 > 1.$$

Ahora bien, si '*P*' fuera definible en función de '*I*', entonces '*P*' tendría que ser el mismo en ambas interpretaciones, puesto que '*I*' lo es. Pero '*P*' no es el mismo y concluimos que '*P*' no puede ser definido en función de '*I*'.

Para aclarar el procedimiento para aplicar el principio de Padoa a cualquier teoría formalizada en lógica de predicados de primer grado con identidad, deseamos ahora hacer más precisa la definición general de independencia de un símbolo primitivo y también caracterizar más precisamente el concepto de que dos interpretaciones de una teoría son diferentes para un símbolo primitivo dado de la mencionada teoría.

Sea **R** un símbolo primitivo de relación de n posiciones de una teoría. Decimos entonces que **R** *depende* de los otros símbolos de la teoría si una fórmula de la forma

$$\mathsf{R}(\mathsf{v}_1, \ldots, \mathsf{v}_n) \leftrightarrow \mathsf{S}$$

puede ser derivada de los axiomas en que (i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son variables distintas, (ii) las únicas variables libres en **S** son $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y (iii) las únicas variables libres en **S** son los otros símbolos primitivos de la teoría.

La estrecha relación entre esta definición de dependencia de un símbolo primitivo de relación y la regla para definir nuevos símbolos de relación en una teoría, es obvia y esperada. Las definiciones de dependencia de símbolos de operación y de constantes individuales son similares y se dejarán como ejercicio.

Deseamos usar ahora la definición de dependencia de los símbolos de relación para precisar la descripción del principio de Padoa en que se basa la prueba de la independencia.[4] Para probar que un símbolo primitivo de relación 'R' de n posiciones es independiente de otros símbolos de una teoría, necesitamos encontrar dos interpretaciones de la teoría, esto es, dos interpretaciones de los axiomas de la teoría tales que

(i) El dominio de ambas interpretaciones es el mismo.

(ii) Las dos interpretaciones son las mismas para todos los otros símbolos primitivos de la teoría.

(iii) Sea 'R_1' la primera interpretación de 'R' y 'R_2' la segunda, entonces 'R_1' y 'R_2' deben ser diferentes en el siguiente respecto: hay elementos $x_1, \ldots, x_n$ en el dominio de interpretación, tales que

(a) '$R_1(x_1, \ldots, x_n)$' es verdadero,

y

(b) '$R_2(x_1, \ldots, x_n)$' es falso.

Para ver que dos representaciones como esas establecen la independencia de 'R', supóngase que R depende de los otros símbolos primitivos de la teoría, esto es, supóngase que hay una fórmula

$$R(x_1, \ldots, x_n) \leftrightarrow \mathrm{S} \tag{1}$$

de la clase exigida por la definición de dependencia, tal que (1) sea deducible de los axiomas de la teoría. Como antes, supongamos que el subíndice '1' se refiere a la primera interpretación y '2' a la segunda interpretación. En ambas interpretaciones debemos tener:

$$R_1(x_1, \ldots, x_n) \leftrightarrow \mathrm{S}_1, \tag{2}$$

y

$$R_2(x_1, \ldots, x_n) \leftrightarrow \mathrm{S}_2, \tag{3}$$

puesto que (1) es una consecuencia lógica de los axiomas de la teoría. Además, puesto que todos los símbolos primitivos, excepto 'R', son los mismos en ambas interpretaciones, tenemos también:

$$\mathrm{S}_1 \leftrightarrow \mathrm{S}_2. \tag{4}$$

---

[4] La exposición original que Padoa hizo de estos temas no es del todo adecuada. Las obras de consulta tipo para una discusión completa son "On the Independence of Undefined Ideas" de J. C. C. McKinsey, en el *Bulletin of the American Mathematical Society*, Vol. 41 (1935) Págs. 291–297, y "Einige methodologische Utersuchungen über die Definierbarkeit der Begriffe", de Alfredo Tarski, en *Erkenntnis*, Vol 5 (1935–1936), Págs. 80 a 100. Una traducción inglesa del artículo de Tarski se encuentra en su obra *Logic, Semantics, Metamathematics*, Oxford, 1956.

De (2), (3) y (4) inferimos que

(5) $$R_1(x_1, \ldots, x_n) \leftrightarrow R_2(x_1, \ldots, x_n),$$

que contradice (a) y (b) de (iii) y prueba que nuestra suposición de dependencia es absurda.

La definición de *alteridad** de interpretación es similar para símbolos de operación y constantes individuales;[5] precisamente la misma clase de argumento como el que se acaba de dar demuestra que encontrar las dos interpretaciones apropiadamente diferentes es adecuado para probar la independencia de los símbolos de operación o de las constantes individuales.

Un ejemplo sencillo ilustrará cómo se usa el principio de Padoa para probar la independencia de un símbolo de operación. Consideramos la teoría cuyos símbolos primitivos son un símbolo 'P' de relación de una posición (esto es, 'P' denota una propiedad) y un símbolo de operación binaria '$\circ$'. Los axiomas de la teoría son

A1. $P(x) \,\&\, P(y) \rightarrow P(x \circ y)$.
A2. $P(x) \,\&\, {-P(y)} \rightarrow -P(x \circ y)$.
A3. $x \circ y = y \circ x$.

Indicamos la primera y segunda interpretaciones por medio de subíndices como anteriormente. El dominio común de interpretación es el conjunto de enteros positivos, y

$$P_1(x) \leftrightarrow P_2(x) \leftrightarrow x \text{ es un entero par}$$
$$x \circ_1 y = x + y,$$
$$x \circ_2 y = x + y + 2.$$

Verificamos fácilmente que ambas interpretaciones satisfacen los axiomas. La independencia de '$\circ$' se sigue del hecho de que tenemos:

$$1 \circ_1 2 = 3$$
$$1 \circ_2 2 = 5 \neq 3.$$

Sin entrar en detalles, puede mencionarse que el principio de Padoa puede extenderse fácilmente a teorías que suponen en su formalización no solamente lógica de predicados de primer orden con identidad, sino también una buena cantidad de matemáticas clásicas. Probablemente algunas de las más interesantes aplicaciones del

* Del vocablo latino "alter» derivamos "alteridad" para significar "diferencia", con la modalidad que el autor le atribuye en el contexto y para la cual crea, en inglés, el neologismo "differentness". (N. del T.)

[5] Una forma clásica de describir esta diferencia, es decir, que el símbolo primitivo cuya independencia se prueba debe tener una *extensión* diferente en ambas interpretaciones.

principio de Padoa a esas teorías "avanzadas" pertenecen al dominio de la ciencia empírica, pues hay un buen volumen de confusa discusión sobre la interdefinibilidad de varios conceptos empíricos. Por ejemplo, en discusiones de los fundamentos de la mecánica, se postula a menudo siguiendo a Ernesto Mach, que el concepto de masa puede ser definido en términos del concepto de aceleración o que el concepto de fuerza puede ser definido en términos de los conceptos de masa y aceleración. Sin embargo, puede demostrarse por aplicación del principio de Padoa que en varias axiomatizaciones plausibles los conceptos de masa y fuerza son cada uno independientes de los otros conceptos primitivos de la mecánica.[6]

EJERCICIOS

Tenga cuidado de no violar la Regla VII §4.2, al hacer estos ejercicios.

1. Considérese la teoría débil de la preferencia discutida en este apartado. Pruébese que el símbolo primitivo de relación '*I*' depende de '*P*'.
2. Defínase la noción de dependencia para símbolos de operación.
3. Caracterícense las dos interpretaciones diferentes necesarias para probar que un símbolo de operación es un símbolo primitivo independiente de una teoría.
4. Considérese la teoría de grupos, tal como se da en el Cap. 5, y basada en tres símbolos primitivos. ¿Es el símbolo de operación '$^{-1}$' del inverso de un elemento de un grupo, dependiente o independiente?
5. Considérense los axiomas para la medición de masa, dados en el Ejercicio 9, § 4.5. Pruébese que los símbolos primitivos '*Q*' y '$\star$' son ambos independientes.
6. Defínase la noción de dependencia para constantes individuales.
7. Caracterícense las dos interpretaciones diferentes necesarias para probar que una constante individual es un símbolo independiente primitivo de una teoría.
8. Considérese la teoría de grupos discutida en el Cap. 5. ¿Es la constante individual '*e*' dependiente o independiente en esta teoría?
9. Considérese la axiomatización de Huntington, para la "parte informal" de *Principia Mathematica*, ya discutida como Ejercicio 12, § 4.1. Los tres símbolos discutidos son un símbolo '*C*' de relación, de una posición, un símbolo de operación binaria '+' y un símbolo de operación unaria '$^{\mathsf{I}}$'; los cinco axiomas son:

A1. *Si* $C(x + y)$, *entonces* $C(y + x)$.
A2. *Si* $C(x)$, *entonces* $C(x + y)$.
A3. *Si* $C(x^{\mathsf{I}})$, entonces $-C(x)$.
A4. *Si* $-C(x^{\mathsf{I}})$, entonces $C(x)$.
A5. *Si* $C(x + y)$ & $C(x^{\mathsf{I}})$, *entonces* $C(y)$.

Para demostrar que el símbolo primitivo '*C*' es independiente, considérese el dominio de interpretación consistente en los números 1, 2, 3, 4. En la primera

[6] Véase §12.5.

interpretación, supóngase que 1 y 2 tienen la propiedad C, esto es, $C_1(1)$ & $C_1(2)$. El símbolo de operación binaria '+' es dado por la siguiente tabla para ambas interpretaciones (el uso de esas tablas finitas es común en la prueba de independencia de axiomas o símbolos primitivos).

+	1	2	3	4
1	1	1	1	1
2	1	2	1	2
3	1	1	3	3
4	1	2	3	4

La tabla se usa de la manera siguiente. Para encontrar qué elemento es $3 + 2$, buscamos la anotación que aparece en la tercera línea y segunda columna y encontramos:

$$3 + 2 = 1.$$

Para la interpretación conjunta de la operación unaria tenemos la tabla siguiente:

x	x^1
1	4
2	3
3	2
4	1

En la segunda interpretación el único cambio está en 'C'. Tenemos ahora: $C_2(1)$ & $C_2(3)$. Se comprueba fácilmente que ambas interpretaciones satisfacen los axiomas. Puesto que

$$C_1(2) \ \& \ -C_2(2),$$

concluimos que 'C' es un símbolo primitivo independiente. Pruébese que los dos símbolos de operación son también independientes.

10. Usando tablas similares a las dadas en el ejercicio anterior, pruébese que los cinco axiomas del Ejercicio 9 son independientes.

PARTE II

TEORIA ELEMENTAL INTUITIVA DE LOS CONJUNTOS

CAPITULO 9

CONJUNTOS

§9.1. Introducción. En éste y en los dos capítulos siguientes desarrollamos la teoría elemental de los conjuntos, de manera intuitiva. El presente capítulo trata de conjuntos arbitrarios; el Cap. 10 de conjuntos que son relaciones, y el Cap. 11 de conjuntos que son funciones.

Puesto que en el Cap. 7 cierta parte de la aritmética se desarrolló conforme a métodos lógicos, a partir de una pequeña lista de axiomas, algunos lectores pueden opinar que sería más adecuado desarrollar la teoría de los conjuntos axiomática y no intuitivamente. Sin embargo, hay motivos poderosos para introducir informalmente los conceptos de teoría de los conjuntos. Todo el mundo está familiarizado con los conceptos de aritmética; una presentación axiomática de la aritmética puede apoyarse en forma continua en hechos que guíen y motiven sus líneas de desarrollo. Aunque los conceptos de la teoría de los conjuntos son lógicamente más sencillos en varios respectos que los de aritmética, no son generalmente familiares. El propósito de los Caps. 9 a 12 es el de suministrar esa familiarización.

§9.2. Membrecía. Por conjunto significamos cualquier tipo de colección de entidades de cualquier clase.[1] Así, podemos hablar del conjunto de todos los americanos, o del conjunto de todos los enteros, o del conjunto de todas las líneas rectas, o del conjunto de todos los círculos que pasan por un punto dado. Como sinónimos de 'conjunto' se usan muchas otras palabras, como 'clase', 'colección' y 'agregado'. Algunas veces usaremos estas palabras en gracia de la variedad literaria.

Decimos de los miembros de un conjunto que *pertenecen* al conjunto; es costumbre usar el símbolo '$\in$', parecido a la letra griega

[1] Aunque el concepto de conjunto fue introducido en el Cap. 2, en conexión con la teoría de la inferencia, la exposición que aquí se presenta es independiente.

epsilón, como abreviatura de 'pertenece a'. Así, escribimos

Isabel II pertenece a la clase de las mujeres,

o simplemente:

(1) Isabel II $\in$ la clase de las mujeres.

En el lenguaje ordinario, (1) se expresaría como:

Isabel II es una mujer.

Así pues, el verbo 'ser' tiene a menudo el significado de membrecía en un conjunto.

Aquí usamos la palabra 'conjunto' en tal forma que un conjunto queda completamente determinado cuando se dan sus miembros; esto es, si A y B son conjuntos que tienen exactamente los mismos miembros, entonces $A = B$. Así escribimos:

El conjunto de triángulos ~~rectángulos~~ equiláteros = el conjunto de triángulos equiángulos,

porque algo pertenece al primer conjunto si y sólo si pertenece al segundo, puesto que un triángulo es equilátero si y sólo si es equiángulo. Este principio general de identidad para conjuntos se llama comúnmente el *principio de extensionalidad para conjuntos;* puede formularse simbólicamente así:

(2) $$A = B \leftrightarrow (x)(x \in A \leftrightarrow x \in B).$$

Algunas veces resulta cómodo hablar de un conjunto cuando no se sabe si este conjunto tiene miembros. Un especialista en genética puede desear hablar sobre el conjunto de mujeres cuyos padres, hermanos y esposos son todos hemofílicos, aun cuando no sepa de un solo caso de una mujer como ésas. Un matemático puede querer hablar sobre mapas que no pueden colorearse con menos de cinco colores, aun cuando no pueda probar que esos mapas existen (la cuestión de si existen tales mapas, es, en realidad, un famoso problema no resuelto de matemáticas). Así pues, es conveniente dar al término 'conjunto' un significado suficientemente amplio para que incluya conjuntos *vacíos*, esto es, conjuntos que carecen de miembros.

De nuestro análisis de las implicaciones en el Cap. 1, resulta claro que si A es un conjunto que carece de miembros, entonces el siguiente enunciado es verdadero, puesto que el antecedente siempre es falso:

(3) $$(x)(x \in A \rightarrow x \in B).$$

Y, recíprocamente, si B está vacío, esto es, carece de miembros, entonces es verdad que

(4) $$(x)(x \in B \rightarrow x \in A).$$

De (2), (3) y (4) concluimos que si dos conjuntos A y B son vacíos, entonces:

$$A = B;$$

es decir, sólo hay un conjunto vacío, pues dados dos conjuntos vacíos, se sigue del principio de extensionalidad para conjuntos que los dos conjuntos son idénticos. De aquí que hablemos del conjunto vacío, que designamos con la letra griega lambda mayúscula:

$$\Lambda.$$

Λ es el conjunto tal que para toda x, x no pertenece a Λ, esto es, simbólicamente

$$(x)-(x \in \Lambda),$$

utilizamos '$x \notin \Lambda$', como expresión abreviada de '$-(x \in \Lambda)$' y escribimos:

$$(x)(x \notin \Lambda).$$

Encontraremos cómodo, en general, usar la notación '$\notin$' para indicar que algo *no pertenece* a un conjunto.

A menudo describiremos un conjunto escribiendo los nombres de sus miembros, separados por comas, y encerrando el total en llaves. Por ejemplo, por

$$\{\text{Roosevelt, Parker}\}$$

significamos el conjunto consistente en los dos candidatos principales de la elección presidencial americana de 1904. Por

$$\{1, 3, 5\}$$

significamos el conjunto consistente en los tres primeros enteros positivos e impares. Está claro que

$$\{1, 3, 5\} = \{1, 5, 3\}$$

(pues ambos conjuntos tienen el mismo número de miembros; el orden en que se asientan los miembros de un conjunto carece de importancia). Además

$$\{1, 1, 3, 5\} = \{1, 3, 5\}$$

(porque no se cuenta dos veces un elemento de un conjunto).

Los miembros de un conjunto pueden a su vez ser conjuntos. Así, un partido político puede ser concebido como un conjunto de

personas y puede ser conveniente hablar del conjunto de partidos políticos de un país dado. De modo similar, podemos tener conjuntos cuyos miembros sean conjuntos enteros; por ejemplo, por medio de

$$\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$$

significamos el conjunto que tiene exactamente tres miembros, a saber, $\{1, 2\}$, $\{3, 4\}$ y $\{5, 6\}$. Por

$$\{\{1, 2\}, \{2, 3\}\}$$

significamos el conjunto cuyos miembros son $\{1, 2\}$ y $\{2, 3\}$. Por

$$\{\{1, 2\}, \{1\}\}$$

significamos el conjunto cuyos dos miembros son los conjuntos $\{1, 2\}$ y $\{1\}$.

Un conjunto que consta de un solo miembro, no debe considerarse idéntico a ese miembro. Así, el conjunto $\{\{1, 2\}\}$ no es idéntico al conjunto $\{1, 2\}$, esto resulta claro del hecho de que $\{1, 2\}$ tiene dos miembros y $\{\{1, 2\}\}$ tiene un miembro a saber, $\{1, 2\}$). De modo semejante,

$$\{\text{Isabel II}\} \neq \text{Isabel II},$$

porque Isabel II es una mujer, mientras que {Isabel II} es un conjunto. Así, el conjunto de sillas no es un miembro del conjunto de sillas: esto es, el conjunto de sillas no es él mismo una silla. Esta observación ilustra la muy grande diferencia entre identidad y membrecía; pues la aserción de que

$$A = A$$

es siempre verdadera, mientras que

$$A \in A$$

es generalmente falsa.[2]

La relación de membrecía difiere también de la relación de identidad en que no es simétrica; de $A \in B$ no se sigue que $B \in A$. Por ejemplo, tenemos

$$2 \in \{1, 2\},$$

pero

$$\{1, 2\} \notin 2.$$

Además, la relación de membrecía no es transitiva: de $A \in B$ y $B \in C$ no se sigue que $A \in C$. Así, por ejemplo, tenemos:

[2] En la mayor parte de los sistemas tipo de teoría axiomática de los conjuntos, ningún conjunto puede ser miembro de sí mismo.

$$2 \in \{1, 2\}$$

y

$$\{1, 2\} \in \{\{1, 2\}, \{3, 4\}\}$$

pero

$$2 \notin \{\{1, 2\}, \{3, 4\}\},$$

pues los únicos miembros de $\{\{1, 2\}, \{3, 4\}\}$ son $\{1, 2\}$ y $\{3, 4\}$, y ninguno de estos conjuntos es igual a 2.

Debe advertirse que si, por ejemplo, $\{a, b\}$ es cualquier conjunto con dos miembros, entonces, para toda x, $x \in \{a, b\}$ si y sólo si $x = a$, o $x = b$, esto es, simbólicamente

$$(x)(x \in \{a, b\} \leftrightarrow (x = a \vee x = b)).$$

De modo semejante, si $\{a, b, c\}$ es un conjunto de tres miembros, entonces $x \in \{a, b, c\}$ si y sólo si $x = a$ o $x = b$ o $x = c$. Por esta razón decimos justamente que $2 \notin \{\{1, 2\}, \{3, 4\}\}$; porque si $x \in \{\{1, 2\}, \{3, 4\}\}$, entonces $x = \{1, 2\}$ o $x = \{3, 4\}$, y puesto que $2 \neq \{1, 2\}$ y $2 \neq \{3, 4\}$, se sigue que $2 \notin \{\{1, 2\}, \{3, 4\}\}$.

Debería notarse también que hay una estrecha relación entre decir que algo tiene una propiedad y decir que pertenece a un conjunto: una cosa tiene una propiedad dada si y sólo si pertenece al conjunto de cosas que tienen la propiedad.[3] Así pues, decir que 6 tiene la propiedad de ser un número par equivale a decir que 6 pertenece al conjunto de los números pares. Puesto que podemos siempre en esta forma expresar las cosas en términos de membrecía en conjuntos, en vez de en términos de la posesión de propiedades, no encontramos necesario hacer una exposición más pormenorizada sobre las propiedades.

En el §5.1 expresamos el principio de la identidad de los indiscernibles en términos de propiedades. Expresado en términos de membrecía, este principio toma la forma: si y pertenece a cada uno de los conjuntos a que x pertenece, entonces $y = x$. Así formulado, el principio tiene quizás un carácter más obvio que cuando fue formulado en función de propiedades, porque $x \in \{x\}$ (esto es, x pertenece al conjunto cuyo único miembro es x) y de aquí que, si y pertenece a cada uno de los conjuntos a que x pertenece, concluimos que $y \in \{x\}$, de manera que $y = x$.

§9.3. Inclusión. Si A y B son conjuntos tales que todo miembro de A es también miembro de B, entonces llamamos a A un *subconjunto* de B, o decimos que A está incluido en B. A menudo

[3] Este principio es llamado algunas veces *axioma de Cantor para los conjuntos*, por el fundador de la teoría de los conjuntos, G. Cantor (1845-1918). Este principio debe ser adecuadamente restringido para evitar contradicciones.

usamos el signo '$\subseteq$' como abreviatura de 'está incluido en'. Así, podemos escribir, por ejemplo:

> El conjunto de los americanos es un subconjunto del conjunto de los hombres,

o:

> El conjunto de los americanos está incluido en el conjunto de los hombres,

o sencillamente:

El conjunto de los americanos $\subseteq$ el conjunto de los hombres. Simbólicamente tenemos:

(1) $$A \subseteq B \leftrightarrow (x)(x \in A \rightarrow x \in B).$$

Está claro que todo conjunto es un subconjunto de sí mismo; esto es, para todo conjunto A, se tiene: $A \subseteq A$. Además, la relación de inclusión es transitiva, esto es, si $A \subseteq B$ y $B \subseteq C$, entonces $A \subseteq C$ (pues si todo miembro de A es un miembro de B, y todo miembro de B es un miembro de C, entonces todo miembro de A es un miembro de C). En cambio, la relación de inclusión no es simétrica; así, $\{1, 2\} \subseteq \{1, 2, 3\}$, pero no es cierto que $\{1, 2, 3\} \subseteq \{1, 2\}$.

Es intuitivamente obvio que identidad, membrecía e inclusión son conceptos distintos y es interesante observar que su distinción puede inferirse simplemente de consideraciones de simetría y transitividad. Así, la inclusión no es lo mismo que la identidad, pues la identidad es simétrica, mientras que la inclusión no lo es. A su vez inclusión no es igual que membrecía, pues la inclusión es transitiva y la membrecía no. Anteriormente hemos visto también que identidad no es lo mismo que membrecía, puesto que la identidad es al mismo tiempo simétrica y transitiva, mientras que la membrecía no tiene ninguna de esas dos características. En el lenguaje cotidiano, los tres conceptos se expresan por medio del recargado verbo 'ser'. Por ejemplo, en el lenguaje cotidiano escribimos:

> Isabel II es la actual reina de Inglaterra,
> Isabel II es una mujer,
> Las mujeres son seres humanos.

Pero en el lenguaje más exacto que aquí se desarrolla, tenemos:

> Isabel II = la actual reina de Inglaterra,
> Isabel II $\in$ la clase de las mujeres,

La clase de las mujeres $\subseteq$ la clase de los seres humanos.

Cuando $A \subseteq B$, no queda excluida la posibilidad de que $A = B$; puede también suceder que $B \subseteq A$, de manera que A y B tengan exactamente el mismo número de miembros y de aquí que sean idénticos.

Cuando $A \subseteq B$, pero $A \neq B$, llamamos a A un *subconjunto propiamente dicho* de B. Usamos '$\subset$' como abreviatura de 'es un subconjunto propiamente dicho de'. Así,

$$\{1, 2\} \subset \{1, 2, 3\}$$

es verdadero, como también lo es

$$\{1, 2\} \subseteq \{1, 2, 3\};$$

pero

$$\{1, 2, 3\} \subset \{1, 2, 3\}$$

es falso, aunque

$$\{1, 2, 3\} \subseteq \{1, 2, 3\}$$

es, por supuesto, verdadero.

Simbólicamente tenemos:

$$A \dot{\subset} B \leftrightarrow A \subseteq B \;\&\; A \neq B.$$

EJERCICIOS

1. ¿Cuáles de los siguientes enunciados son verdaderos (para todos los conjuntos A, B y C)?

(a) Si $A = B$ y $B = C$, entonces $A = C$.
(b) Si $A \in B$ y $B \in C$, entonces $A \in C$.
(c) Si $A \subseteq B$ y $B \subseteq C$, entonces $A \subseteq C$.
(d) Si $A = B$ y $B \in C$, entonces $A \in C$.
(e) Si $A \in B$ y $B = C$, entonces $A \in C$.
(f) Si $A \in B$ y $B \subseteq C$, entonces $A \in C$.
(g) Si $A \subseteq B$ y $B \in C$, entonces $A \in C$.
(h) Si $A \subset B$ y $B \in C$, entonces $B \subset C$.
(i) Si $A \subset B$ y $B \subseteq C$, entonces $A \subset C$.
(j) Si $A \in B$ y $B \subset C$, entonces $A \in C$.
(k) Si $A \in B$ y $B \subset C$, entonces $A \in C$.

2. Por cada uno de los enunciados del Ejercicio 1 que sea falso dése un ejemplo de conjuntos particulares A, B y C que demuestren que el enunciado no es en general verdadero.

3. Dese un ejemplo de conjuntos A, B, C, D que satisfagan las condiciones

$$A \subset B$$
$$B \in C$$
$$C \subset D$$
$$D = E$$

4. ¿En qué consiste el error del siguiente argumento?

> Sócrates es un hombre. Los hombres son numerosos. Sócrates es numeroso.

5. ¿En qué consiste el error del siguiente argumento?

> Los gatos machos son gatos. Los gatos son una especie. En consecuencia, los gatos machos son una especie.

6. En cada uno de los siguientes ejemplos decídase cuál de los siguientes enunciados son verdaderos: $A \in B$, $A \subseteq B$, $A \subset B$, $A = B$.

EJEMPLO (A).

$A = \{1, \{1\}, \text{Roosevelt}, 4\}$,
$B = \{1, \{1\}, \text{Roosevelt}, \text{Churchill}\}$.

EJEMPLO (B).

$A =$ el conjunto de números positivos.
$B =$ el conjunto de enteros positivos y negativos.

EJEMPLO (C).

A es el conjunto que consiste de lo siguiente: el número 5, el conjunto consistente en Roosevelt, el conjunto consistente en el conjunto consistente en el número 1.

B es el conjunto consistente en lo siguiente: el número 5, Roosevelt, el conjunto consistente en el conjunto consistente en Roosevelt, el número 1.

EJEMPLO (D).

$A = \{1, 3, 4, 2, 9\}$
$B = \{1, 2 + 1, 1 + 8, 10, 2 + 0, 1008, 4\}$

EJEMPLO (E).

$A =$ el conjunto cuyos miembros son los siguientes: el conjunto de todos los presidentes de los Estados Unidos en el siglo diecinueve, Truman, McKinley, el número positivo más pequeño divisible entre 5.

$B =$ el conjunto cuyos miembros son los siguientes: el conjunto de todos los presidentes de los Estados Unidos hasta 1953, el número 10.

7. ¿Cuáles de los siguientes enunciados son verdaderos para todos los conjuntos?

(a) $A \not\subset B \;\&\; B \not\subset C \rightarrow A \not\subset C$
(b) $A \neq B \;\&\; B \neq C \rightarrow A \neq C$
(c) $A \in B \;\&\; -(B \subseteq C) \rightarrow A \not\subset C$
(d) $A \subset B \;\&\; B \subseteq C \rightarrow -(C \subset A)$
(e) $A \subseteq B \;\&\; B \in C \rightarrow A \not\subset C$

§9.4. El conjunto vacío. Como se indicó anteriormente, el conjunto vacío se caracteriza por la propiedad de que, para toda x, $x \notin \Lambda$.

Debe advertirse que, aunque nada pertenece al conjunto vacío, éste puede muy bien ser miembro de otro conjunto. Así, si hablamos del conjunto de todos los subconjuntos del conjunto $\{1, 2\}$, estamos hablando del conjunto $\{\{1, 2\}, \{1\}, \{2\}, \Lambda\}$, que tiene cuatro miembros; el conjunto de tres miembros $\{\{1, 2\}, \{1\}, \{2\}\}$, por otra parte, es el conjunto de todos los subconjuntos no vacíos de $\{1, 2\}$.

Recordamos el hecho de que un conjunto A es un subconjunto de un conjunto B si y sólo si todo miembro de A es también un miembro de B, esto es, si y sólo si, para todo x, si $x \in A$, entonces $x \in B$. En particular, el conjunto vacío Λ es un subconjunto de un conjunto B si y sólo si, para toda x, si $x \in \Lambda$, entonces $x \in B$. Sin embargo, puesto que siempre $x \notin \Lambda$, siempre es verdadero que si $x \in \Lambda$, entonces $x \in B$. Así pues, para todo conjunto B tenemos

$$\Lambda \subseteq B.$$

Esto es, el conjunto vacío es un subconjunto de todos los conjuntos. Además, el conjunto vacío es el único conjunto que es un subconjunto real del conjunto vacío; pues si $B \subseteq \Lambda$, entonces, puesto que tenemos también $\Lambda \subseteq B$, podemos también concluir que $B = \Lambda$.

Podemos resumir estos hechos sobre el conjunto vacío como sigue:

(i) $$(x)(x \notin \Lambda),$$

(ii) $$(\exists A)(\Lambda \in A) \ \& \ (\exists A)(\Lambda \notin A),$$

(iii) $$(A)(\Lambda \subseteq A),$$

(iv) $$(A)(A \subseteq \Lambda \leftrightarrow A = \Lambda).$$

§9.5. Operaciones con conjuntos. Si A y B son conjuntos, entonces, por la intersección de A y B (en símbolos $A \cap B$), significamos el conjunto de todas las cosas que pertenecen tanto a A como a B. Así, para toda x, $x \in (A \cap B)$, si y sólo si $x \in A$ y $x \in B$; esto es, simbólicamente

(1) $$(x)(x \in A \cap B \leftrightarrow x \in A \ \& \ x \in B).$$

Si A es el conjunto de todos los americanos y B es el conjunto de todas las personas de ojos azules, entonces $A \cap B$ es el conjunto de todos los americanos de ojos azules.

Si A es el conjunto de todos los hombres y B es el conjunto de todos los animales que pesan más de diez toneladas, entonces $A \cap B$ es el conjunto de todos los hombres que pesan más de diez toneladas. En este caso notamos que $A \cap B$ es el conjunto vacío (a pesar del hecho de que $A \neq \Lambda$ y $B \neq \Lambda$, puesto que algunas ballenas pesan

más de diez toneladas). Cuando $A \cap B = \Lambda$, decimos que A y B son *mutuamente exclusivos*.

El significado que damos al término 'intersección' es semejante al que se le da en geometría elemental, donde por intersección de dos círculos, por ejemplo, significamos los puntos que quedan en ambos círculos. Algunos autores usan, en vez de '$\cap$', el punto '.' que se usa en álgebra para la multiplicación; esos autores a menudo hablan del "producto" de dos conjuntos, en vez de su intersección.

Si A y B son conjuntos, entonces por la unión de A y B (en símbolos, $A \cup B$) significamos el conjunto de todas las cosas que pertenecen cuando menos a uno de los conjuntos A y B. Así, para toda x, $x \in (A \cup B)$, si y sólo si o $x \in A$, o $x \in B$. (Nótese que, como se explicó en el Cap. 1, usamos el conectivo 'o' en su sentido *no exclusivo*; '$x \in A$ o $x \in B$' es falso solamente cuando se dan ambos casos: $x \notin A$ y $x \notin B$). Simbólicamente

(2) $$(x)(x \in A \cup B \leftrightarrow x \in A \vee x \in B).$$

Si A es el conjunto de todos los animales y B es el conjunto de todas las plantas, entonces $A \cup B$ es el conjunto de todos los organismos vivientes. A menudo se desea considerar la unión de dos conjuntos, aun cuando no son mutuamente exclusivos. Por ejemplo, si A es el conjunto de todos los seres humanos adultos y B es el conjunto de todas las personas menores de 40 años, entonces $A \cup B$ es el conjunto de todos los seres humanos.

Algunos autores usan el signo de adición '+', en vez de '$\cup$', y llaman a la unión de dos conjuntos su "suma".

Si A y B son dos conjuntos, entonces por *diferencia* de A y B (en símbolos, $A \sim B$) significamos el conjunto de todas las cosas que pertenecen a A pero no a B. Así, para toda x, $x \in A \sim B$, si y sólo si $x \in A$ y $x \notin B$; esto es, simbólicamente:

(3) $$(x)(x \in A \sim B \leftrightarrow x \in A \,\&\, x \notin B).$$

Si A es el conjunto de todos los seres humanos y B es el conjunto de todos los humanos femeninos, entonces $A \sim B$ es el conjunto de todos los humanos masculinos. A menudo se desea considerar la diferencia de dos conjuntos A y B, aun cuando B no sea un subconjunto de A. Por ejemplo, si A es el conjunto de seres humanos y B es el conjunto de todos los animales femeninos, entonces $A \sim B$ sigue siendo el conjunto de todos los humanos masculinos y $B \sim A$ es el conjunto de todos los animales femeninos que pertenecen a especies no humanas.

Estas operaciones sobre conjuntos (intersección, unión y diferencia) pueden, por supuesto, ser repetidas. Así, supóngase, por ejemplo, que

$$A = \{1, 2\},$$
$$B = \{1, 3, 5\},$$
$$C = \{2, 3, 5, 7\},$$
$$D = \{4, 5, 6, 7\};$$

entonces

$$A \cup B = \{1, 2\} \cup \{1, 3, 5\} = \{1, 2, 3, 5\}$$

y de aquí que

$$C \cap (A \cup B) = \{2, 3, 5, 7\} \cap \{1, 2, 3, 5\} = \{2, 3, 5\}$$

y de aquí que

$$D \sim [C \cap (A \cup B)] = \{4, 5, 6, 7\} \sim \{2, 3, 5\} = \{4, 6, 7\}.$$

De modo semejante, puesto que

$$C \cup D = \{2, 3, 5, 7\} \cup \{4, 5, 6, 7\} = \{2, 3, 4, 5, 6, 7\},$$

tenemos:

$$(A \cup B) \cap (C \cup D) = \{1, 2, 3, 5\} \cap \{2, 3, 4, 5, 6, 7\} = \{2, 3, 5\}.$$

EJERCICIOS

1. Si A es el conjunto de todos los enteros positivos, y B es el conjunto de todos los enteros mayores que 10, ¿qué son los siguientes?

$$A \cup B \qquad A \sim B$$
$$A \cap B \qquad B \sim A$$

2. Si A es un conjunto cualquiera, ¿qué son los siguientes?

$$A \cap \Lambda \qquad A \sim \Lambda$$
$$A \cup \Lambda \qquad \Lambda \sim A$$

3. Haciendo:

A = el conjunto de enteros positivos $\qquad C = \{2, 4\}$

$B = \{3, 5\} \qquad D = \{1, 2\}$

encuentre lo siguiente:

$$A \sim B \qquad (B \cup C) \cap (B \cup D)$$
$$A \sim C \qquad A \sim (C \cap D)$$
$$A \sim D \qquad (A \sim C) \cup (A \sim D)$$

4. Haciendo:

$$A = \{1, 2\} \qquad F = \{\{1, 3\}, 1, 2\}$$
$$B = \{\{3, 4\}, 1, 7\} \qquad G = \{\{1, 2\}, 1, 7\}$$
$$C = \{\{3, 4\}, 1, 2\} \qquad H = \{\{1, 2\}, 1, 2\}$$
$$D = \{\{1, 3\}, 4, 7\} \qquad I = \{\{1, 2\}, \{1\}, \{2\}\}$$
$$E = \{\{1, 3\}, 1, 7\} \qquad J = \{\{1\}, \{2\}\}$$

encuentre lo siguiente:

$A \cap B$ $\quad$ $A \cap G$
$A \cap C$ $\quad$ $A \cap H$
$A \cap D$ $\quad$ $A \cap I$
$A \cap E$ $\quad$ $A \cap J$
$A \cap F$

5. Encuentre lo siguiente:

$\Lambda \cap \{\Lambda\}$
$\{\Lambda\} \cap \{\Lambda\}$
$\{\Lambda, \{\Lambda\}\} \sim \Lambda$
$\{\Lambda, \{\Lambda\}\} \sim \{\Lambda\}$
$\{\Lambda, \{\Lambda\}\} \sim \{\{\Lambda\}\}$

6. Haciendo:

$A = \{1\}$ $\quad$ $C = \{1, 2\}$ $\quad$ $E = \{1, \{1, \{1\}\}\}$
$B = \{1, \{1\}\}$ $\quad$ $D = \{1, 2, \{1\}\}$

encuentre lo siguiente:

$A \cap B$ $\quad$ $(C \cup D) \sim B$
$A \cup B$ $\quad$ $(A \cap D) \sim E$
$(A \cup B) \cap C$ $\quad$ $\{B\} \cap E$
$\{A\} \cap B$ $\quad$ $(\{A\} \cup D) \cap (E \sim C)$.

7. Usando los conjuntos del Ejercicio 6, ¿cuál de los siguientes enunciados es verdadero?

(a) $A \in B$
(b) $A \subseteq B$
(c) $B \in E$
(d) $B \subseteq E$
(e) $C \in D$
(f) $C \subseteq D$
(g) $B \subset D$
(h) $B \sim A \in D$
(i) $E \sim B \subseteq A$

8. Tomando el símbolo de relación binaria '$\in$' como el único símbolo primitivo de la teoría de los conjuntos, defina los símbolos de operación de intersección, unión y diferencia en una forma que satisfaga las reglas de definición para símbolos de operación dadas en el Cap. 8.

§9.6. Dominios de individuos. A menudo se está interesado, no en todos los conjuntos posibles, sino únicamente en todos los subconjuntos de algún conjunto fijo. Así, en sociología, por ejemplo, podría resultar natural hablar mayormente de conjuntos de seres humanos, y hablar en la inteligencia de que cuando se mencionara un conjunto, habría de tomarse como un conjunto de personas, a menos de que se hiciera una explícita indicación en contrario. En una exposición de esa naturaleza se podría decir, por ejemplo, 'el conjunto de albinos' y se entendería que se estaba haciendo referencia únicamente al conjunto de personas albinas, y no también a monos, ratones y otros animales albinos. De modo semejante, en una exposición geométrica es natural usar la palabra conjunto para

significar un 'conjunto de puntos'. (Algunas veces en matemáticas se hace tomar un sentido especial a algunas de las diversas palabras que anteriormente se ha dicho que aquí se toman como sinónimos de 'conjunto'; un geómetra podría, por ejemplo, adoptar la convención de hablar de *conjuntos* de puntos, *clases* de conjuntos de puntos y *agregados* —o quizás *familias*— de curvas geométricas.)

Cuando un conjunto fijo D se toma como dado en este sentido, y la discusión se limita a subconjuntos de D, llamamos a D el *dominio de individuos*, o algunas veces el *dominio del discurso*. Así, el dominio de individuos de la discusión sociológica mencionada anteriormente es el conjunto de todos los seres humanos.

Denotaremos el dominio de individuos por 'V'. Es importante recordar que aunque 'Λ', como se ha visto anteriormente, representa una entidad determinada unívocamente (el conjunto vacío), el símbolo 'V' se interpreta de modo diferente en discusiones diferentes. En un contexto 'V' puede representar el conjunto de seres humanos, en otro el conjunto de puntos del espacio y en otro el conjunto de enteros positivos.

Cuando se trata de un dominio fijo de individuos V, es conveniente introducir un símbolo especial para la diferencia de V y un conjunto A:

$$\sim A = \text{V} \sim A.$$

Llamamos a $\sim A$ el *complemento de A*. Más generalmente, la diferencia $B \sim A$ de B y A se llama el *complemento de A con relación AB*; de manera que el complemento de un conjunto es simplemente su complemento con relación al dominio dado de individuos.

Como era de esperarse, podemos encontrar el complemento de un conjunto que a su vez resulta de operaciones con otros conjuntos. Por ejemplo, sea

$$\text{V} = \{1, 2, 3\}$$
$$A = \{1, 2\}$$
$$B = \{2, 3\}.$$

Entonces

$$\sim A = \text{V} \sim A = \{1, 2, 3\} \sim \{1, 2\} = \{3\},$$

y, correspondientemente,

$$\sim(A \cup B) = \text{V} \sim (A \cup B) = \{1, 2, 3\} \sim (\{1, 2\} \cup \{2, 3\})$$
$$= \{1, 2, 3\} \sim \{1, 2, 3\} = \Lambda.$$

En el §9.9 se mencionarán algunos otros hechos sobre la operación de complementación.

EJERCICIOS

1. Si

$$V = \{1, 2, 3, 4, 5\} \qquad A = \{1, 2\} \quad B = \{2, 3\}$$

qué son las siguientes:

$\sim A$	$\sim A \cap \sim B$	$A \sim (\sim B)$
$\sim B$	$\sim (A \cap B)$	$\sim A \sim B$
		$\sim A \sim (\sim B)$

2. Sea V el conjunto de todos los enteros positivos y sean

A = conjunto de todos los enteros positivos,
B = conjunto de todos los enteros positivos impares,
C = conjunto de todos los enteros positivos mayores que 10,
D = conjunto de todos los enteros positivos menores que 15.

Encuéntrese:

(a) $\sim A$	(e) $D \sim C$	(i) $A \sim (\sim C)$
(b) $\sim (A \cup B)$	(f) $\sim (D \sim C)$	(j) $(A \cap D) \sim (\sim B)$
(c) $\sim (A \cap B)$	(g) $C \cup \sim D$	(k) $A \sim (\sim C \cup D)$
(d) $\sim C$	(h) $\sim C \cap D$	

3. Considérense los axiomas para grupos dados en el Cap. 5. Sea el dominio de interpretación el conjunto de todos los subconjuntos de algún dominio dado V de individuos. Interprétase como complementación el símbolo de operación inversa. Si el símbolo de operación binaria es interpretado como:

(a) intersección de conjuntos
(b) unión de conjuntos
(c) diferencia de conjuntos

¿quedan satisfechos los tres axiomas sobre grupos? Si no, indíquese qué axiomas no se satisfacen y dé contraejemplos explícitos.

4. Igual que el Ejercicio 3, con la excepción de que el elemento de identidad es interpretado como el dominio V de individuos.

§9.7. Transcripción del lenguaje cotidiano.[4] Esta sección está dedicada al problema de transcribir oraciones del lenguaje cotidiano, al simbolismo que se ha estado desarrollando en este capítulo. Debe advertirse, como se indicó en el Cap. 3, que el significado del lenguaje cotidiano no es tan uniforme que puedan darse reglas de transcripción categóricas y exentas de ambigüedad. En el lenguaje cotidiano usamos a menudo la misma palabra para conceptos esencialmente diferentes (por ejemplo, 'es', tanto para '$\in$' como para '$\subseteq$'; y, algunas veces por elegancia literaria, usamos a menudo palabras diferentes para el mismo concepto ('es', 'es un subconjunto de' y 'está incluido en', por ejemplo, para '$\subseteq$').

[4] Esta sección puede omitirse sin pérdida de continuidad.

Consideramos aquí solamente las oraciones que pueden transcribirse a un simbolismo consistente solamente en letras que representan conjuntos, en paréntesis, y en los simbolismos siguientes:

$$\cap, \cup, \sim, \Lambda, =, \neq, \subseteq.$$

Ese simbolismo puede manejar muy bien enunciados que impliquen predicados de una posición, pero no es adecuado para las complejidades de pensamiento que pueden expresarse por medio de predicados de varias posiciones. Este simbolismo es esencialmente equivalente a la teoría clásica del silogismo. Es importante hacer notar que no estamos usando aquí la noción de membrecía; nos limitamos a conjuntos que están todos al mismo nivel: subconjuntos de algún dominio fijo de individuos.

Un enunciado lingüístico de la forma 'todos los... son...', donde los dos huecos pueden ser llenados con nombres comunes como 'hombres' o 'americanos' o 'filósofos', significa, por supuesto, que el conjunto de cosas descrito por el primer nombre es un subconjunto del conjunto dę cosas descrito por el segundo nombre. Así, por ejemplo,

(1) Todos los americanos son filósofos

significa

El conjunto de americanos $\subseteq$ el conjunto de filósofos,

o, usando '*A*' como abreviatura de "el conjunto de americanos", y '*P*' como abreviatura de "el conjunto de filósofos":

$$A \subseteq P.$$

Podemos también expresar el significado de este enunciado en otras formas equivalentes:

$$A \cup P = P,$$

o:

$$A \cap \sim P = \Lambda,$$

etc., y estos modos de expresión resultan a menudo útiles. En estas discusiones se fija fácilmente un dominio de individuos y, de este modo, la magnitud de conjuntos complementarios como $\sim P$. Aquí, por ejemplo, V puede tomarse como el conjunto de todos los seres humanos.

Usamos el mismo modo de transcripción de enunciados de la forma "todos los .... son ..." también cuando el segundo hueco se llena con un adjetivo. Por ejemplo, tomamos:

(2) Todos los americanos son mortales

para significar:

La clase de los americanos $\subseteq$ la clase de los seres mortales,

o, usando abreviaturas obvias:

$$A \subseteq M.$$

Algunas veces, su cambio, en contextos de este tipo, se suprime la palabra "todos", escribiendo, por ejemplo:

Los tiranos son mortales

en vez de:

(3) Todos los tiranos son mortales,

o:

Las mujeres son volubles,

en vez de:

(4) Todas las mujeres son volubles,

que transcribiríamos, respectivamente, por:

$$T \subseteq M$$

y:

$$W \subseteq F.$$

Es necesario estar en guardia contra la transcripción de enunciados de esta clase, pues el lenguaje ordinario usa la misma forma para expresar ideas esencialmente diferentes. Así, como hemos visto anteriormente:

(5) Los hombres son numerosos

no significa:

El conjunto de los hombres $\subseteq$ el conjunto de cosas numerosas

(esto es, que cada hombre sea numeroso), sino más bien, suponiendo que M sea el conjunto de los hombres y N el conjunto de los conjuntos que tienen numerosos miembros:

$$M \in N.$$

De modo semejante:

(6) Los apóstoles son doce

significa que el conjunto de los apóstoles pertenece al conjunto de los conjuntos que tienen solamente doce miembros.

Correspondiendo a la distinción que acabamos de hacer entre membrecía e inclusión, la antigua lógica hacía una distinción entre la aplicación "distributiva" y la "colectiva" del predicado al sujeto. Usando esta terminología, se dice que en (1), (2), (3) y (4) el predicado se aplica distributivamente al sujeto y que en (5) y (6) se aplica colectivamente.

Un enunciado idiomático de la forma "algunos ... son ...", donde los huecos se llenan con nombres comunes (o quizás el segundo hueco se llena con un adjetivo) significa que existe algo que se describe por medio de ambos términos, esto es, que no está vacía la intersección de los dos conjuntos correspondientes. Así por ejemplo,

(7) Algunos americanos son filósofos

significa que existe cuando menos una persona que es tanto americano como filósofo y, en consecuencia, se transcribe como:

$$A \cap P \neq \Lambda.$$

Aunque un enunciado de la forma de (7) implica que no están vacíos los conjuntos correspondientes al sujeto y al predicado, no puede extraerse esa inferencia de un enunciado de la forma de (1). Así, por ejemplo, es verdad que

Todos los hombres de tres cabezas y seis ojos,
son hombres de tres cabezas,

pero no es verdad que

Algunos hombres de tres cabezas y seis ojos
son hombres de tres cabezas.

(la lógica moderna difiere en este punto de la antigua, que permitía inferir "Algunos S son P", de "Todos los S son P".)

Un enunciado idiomático de la forma "No ... son ..." (donde, como antes, los huecos se llenan con nombres comunes) significa que nada pertenece simultáneamente al conjunto correspondiente al primer nombre y al conjunto correspondiente al segundo nombre, esto es, que la intersección de ambos conjuntos está vacía. Por ejemplo, la oración:

(8) Ningún americano es filósofo

se transcribe como:

$$A \cap P = \Lambda.$$

Y así, (2) tiene el mismo significado que:

Ningún americano es inmortal

puesto que ambos pueden transcribirse como:

$$A \cap P \sim M = \Lambda.$$

Un enunciado idiomático de la forma "Algunos ... no son ..." (cuando los huecos se llenan con nombres comunes) significa que existe algo que pertenece al conjunto correspondiente al primer nombre y que no pertenece al conjunto correspondiente al segundo nombre, esto es, que la intersección del primer conjunto con el complemento del segundo no está vacía. La oración:

Algunos americanos no son filósofos

se transcribe como:

$$A \cap \sim P \neq \Lambda.$$

Pasamos ahora al problema de transcribir algunos enunciados de tipo más complicado.

La palabra "y" a menudo corresponde a la intersección de conjuntos. Así,

Todos los americanos son limpios y fuertes

se transcribe (usando abreviaturas obvias):

$$A \subseteq C \cap S.$$

Lo mismo se aplica a la palabra "pero"; así, se tiene:

Los novatos son ignorantes, pero entusiásticos

se transcribe:

$$F \subseteq I \cap E.$$

La situación es enteramente diferentes, en cambio, cuando el "y" ocurre en el sujeto y no en el predicado. Así,

(9) Los tontos y los borrachos dicen la verdad

no se transcribe por:

(10) $(F \cap D) \subseteq T$

sino por:

(11) $(F \cup D) \subseteq T;$

pues (9) significa que son verdaderas las dos aseveraciones siguientes:

(12) Todos los tontos dicen la verdad

y:

(13) Todos los borrachos dicen la verdad

y (12) y (13) se transcriben, respectivamente, por:

(14) $$F \subseteq T$$

y:

(15) $$D \subseteq T;$$

y (14) y (15) equivalen juntas a (11). (Debe notarse que (10) dice menos que (11), pues

$$F \cap D \subseteq F \cup D$$

es verdadero para todo F y D —y de aquí que (10) sea verdadero cada vez que (11) sea verdadero—, mientras que una aseveración de la forma (10) puede ser verdadera aun cuando el correspondiente enunciado de la forma (11) sea falsa.)

A menudo el enunciado que ha de transcribirse no contiene ninguna forma del verbo "ser". Así, la aserción:

Algunos franceses beben vino

puede transcribirse por:

$$F \cap W \neq \Lambda,$$

si consideramos a "F" como una representación de franceses y a W como de bebedores de vino.

La aserción:

Algunos americanos beben café y también leche

puede transcribirse por:

$$A \cap C \cap M \neq \Lambda,$$

donde "A" representa el conjunto de los americanos, "C" el conjunto de las personas que beben café y "M" el conjunto de las personas que beben leche. (Aquí hemos adoptado la práctica, que se sigue con frecuencia, de suprimir paréntesis al representar la intersección de dos o más conjuntos, escribiendo simplemente $A \cap C \cap M$, en vez de $A \cap (C \cap M)$; algunas veces adoptaremos una práctica similar con respecto de la representación de la unión de tres o más conjunto. Discutiremos ulteriormente este tema en el § 9.9.)

Todavía son posibles ejemplos aún más complicados. Considérese:

(16) Algunos americanos que toman té, no toman ni café ni leche.

La forma general de este enunciado es: Ciertas S no son P. El sujeto se traduce:

$$A \cap T,$$

donde T = al grupo de bebedores de té, y el predicado se traduce:

$$C \cup M.$$

El enunciado completo (16) entonces se traduce

$$(A \cap T) \cap \sim(C \cup M) \neq \Lambda,$$

lo que también equivale a:

$$A \cap T \cap \sim C \cap \sim M \neq \Lambda,$$

ya que, correspondiendo a las leyes de De Morgan para los conectivos sentenciales, tenemos:

$$(17) \qquad \sim(C \cup M) = \sim C \cap \sim M.$$

Identidades como (17) se estudian con más detalle en § 9.9.

EJERCICIOS

1. Supóngase que un psicólogo desarrolla un experimento de aprendizaje con un grupo de ratas, en un laberinto en T. Sobre la base de una teoría que ha desarrollado, espera ciertas respuestas, con ciertas características. Por motivos de simplificación, consideremos estas características como izquierda o derecha (que corresponden a una vuelta a la izquierda o a la derecha en el laberinto) y reforzada o no reforzada. Para una comparación exacta de lo predicho por una teoría con lo que se observa experimentalmente, distingue entre el conjunto de respuestas predichas y el conjunto de respuestas observadas. Haciendo

V = el conjunto de todas las respuestas posibles,
P = el conjunto de todas las respuestas predichas,
O = el conjunto de todas las respuestas observadas,
L = el conjunto de todas las respuestas izquierdas,
R = el conjunto de todas las respuestas izquierdas,
E = el conjunto de todas las respuestas reforzadas,
U = el conjunto de todas las respuestas no reforzadas,

transcríbase a forma simbólica las siguientes aserciones que podría hacer el psicólogo sobre su experimento.

(a) Se observaron todas las respuestas predichas.
(b) Se predijeron algunas respuestas izquierdas.
(c) Fueron reforzadas todas las preguntas izquierdas predichas.
(d) No se dejó de predecir ninguna respuesta observada.
(e) No se dejó de predecir ninguna respuesta derecha observada.
(f) No se predijeron algunas de las respuestas izquierdas que se observaron.
(g) No se observó ni se predijo ninguna respuesta derecha.
(h) Algunas respuestas derechas, que fueron reforzadas, se observan que no fueron predichas.

(i) Fueron reforzadas algunas preguntas izquierdas que no fueron predichas.
(j) Se observaron todas las respuestas izquierdas predichas y se encontró que eran reforzadas.

2. Haciendo

V = el conjunto de todas las personas,
A = el conjunto de todos los americanos,
C = el conjunto de todas las personas que beben café,
F = el conjunto de todos los franceses,
M = el conjunto de todos los asesinos,
P = el conjunto de todos los filósofos,
T = el conjunto de todas las personas que beben té,
W = el conjunto de todas las personas que beben vino,

transcriba a la forma simbólica los siguientes enunciados:

(a) Algunos americanos que beben vino son filósofos;
(b) Ningún francés es americano;
(c) Personas que beben vino y café beben también té;
(d) Todos los asesinos franceses beben café, té y vino;
(e) Algunos asesinos americanos beben café y té, pero no vino;
(f) Algunos asesinos franceses que beben vino, no beben ni café ni té;
(g) Un filósofo no bebe té ni café;
(h) Algunos franceses son o filósofos o asesinos;
(i) Todos los bebedores de café beben té o vino.

§ 9.8. **Diagramas de Venn.**[5] Al estudiar conjuntos y las relaciones existentes entre los mismos, es útil algunas veces representar diagramáticamente los conjuntos; se traza un rectángulo para representar el dominio de individuos y a continuación círculos u otras figuras dentro del rectángulo y se considera a los puntos situados dentro de los conjuntos que se representan por medio de figuras.[6] Así, por ejemplo, si sabemos de dos conjuntos A y B que son mutuamente exclusivos, podemos representar esta situación por medio del siguiente diagrama:

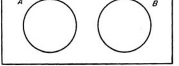


FIG. 1

[5] Esta sección puede omitirse sin pérdida de continuidad.
[6] La idea básica de usar círculos en esta forma se debió al matemático suizo del siglo XVIII Euler. Algunos de los refinamientos que se explicarán más adelante se deben al lógico inglés del siglo XIX Venn. Los diagramas se llaman algunas veces Diagramas de Euler o Diagramas de Venn.

Si sabemos que $A \subseteq B$, podemos representar esta situación por medio de la Fig. 2.

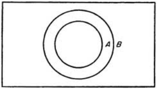


FIG. 2

Una forma más tradicional de describir la Fig. 2, es decir, que todas las *A* son *B*, esto es, todos los miembros de *A* son miembros de *B*. Para reforzar la discusión del apartado anterior, colocaremos entre paréntesis esos enunciados tradicionales, cuando sea conveniente.

Si sabemos de tres conjuntos *A*, *B* y *C*, que $A \subseteq B$ (todas las *A* son *B*) y $B \cap C = \Lambda$ (esto es, ninguna *B* es *C*), entonces podemos representar esta situación por medio de la Fig. 3.

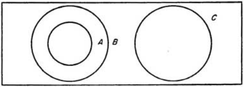


FIG. 3

De la figura se sigue claramente que en este caso debemos también tener $A \cap C = \Lambda$ (ninguna *A* es *C*).

Algunas veces, en vez de tratar de incorporar al diagrama la información dada, trazando simplemente los círculos en la forma adecuada, es conveniente trazar las figuras en forma un tanto arbitraria (de manera que dividan el interior del rectángulo en un número máximo de partes) y a continuación incorporar la información a la figura por otros métodos, como el sombreado de las áreas. Habiendo decidido, digamos, indicar por medio de un *sombreado horizontal*, que una área corresponde al conjunto vacío, indicamos

qua $A \cap B = \Lambda$, por medio de la Fig. 4; que $A \subseteq B$ por medio de la Fig. 5, pues decir que A es un subconjunto de B significa que ninguna parte de A queda fuera de B.

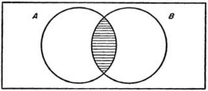


FIG. 4

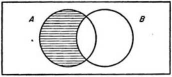


FIG. 5

Usando esos diagramas es fácil a menudo ver qué conclusiones pueden extraerse de los datos obtenidos sobre dos o más conjuntos. Supóngase, por ejemplo, que de los dos conjuntos A y B se dan tanto $A \cap B = \Lambda$, como $A \subseteq B$. La primera aserción (como se indica en la Fig. 4), significa que deben ir sombreadas las partes comunes de A y B, y la segunda (como se indica en la Fig. 5), significa que la parte de A que debe sombrearse es la parte de A que está fuera de B. Obtenemos así la Fig. 6, donde notamos que toda A está sombreada. Vemos de este modo que, juntamente, las dos aseveraciones dadas implican que A es el conjunto vacío.

El sombreado horizontal se usa para indicar la vacuidad de una región. Se necesita otra clase de símbolo para indicar *no vacuidad*,

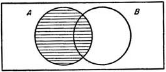


FIG. 6

o *plenitud*. Usaremos un dispositivo de *cruces eslabonadas*.[7] Así, si $A \cap B \neq \Lambda$ (algunas A son B) representamos esta situación por la Fig. 7; la cruz indica que la región común a A y a B no está vacía. Representamos la situación más complicada $A \cap (B \cup C) \neq \Lambda$ (algunas A son B o C), por medio de la Fig. 8.

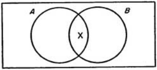


Fig. 7

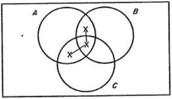


Fig. 8

Las tres cruces de la Fig. 8 están *eslabonadas* para mostrar que cuando menos una de las tres pequeñas regiones no está vacía. Si el eslabonamiento hubiera sido omitido en la Fig. 8, esta figura representaría bastante más que el hecho de que $A \cap (B \cup C) \neq \Lambda$. Si se omitiera el eslabonamiento, podríamos inferir:

(1)	$(A \cap B) \sim C \neq \Lambda$	(la cruz superior)
(2)	$A \cap (B \cap C) \neq \Lambda$	(la cruz media)
(3)	$(A \cap C) \sim B \neq \Lambda$	(la cruz inferior)

Obviamente, cualquiera de las aserciones (1) a (3) implica que $A \cap (B \cup C) \neq \Lambda$, y más. Sin el eslabonamiento, la Fig. 8 diría demasiado.

La situación descrita por

$A \cup B \neq \Lambda$	(Algo es A o B)
$A \cup \sim C \neq \Lambda$	(Algo es A o no C)

[7] Este dispositivo se debe al Prof. Robert McNaughton.

es representada por la Fig. 9. Nótense los dos eslabonamientos separados, uno para cada uno de los dos enunciados existenciales.

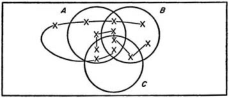


FIG. 9

¿Qué interpretación debe darse a un diagrama en que una cruz y un sombreado ocurren en la misma región? Supóngase, por ejemplo, que tenemos

(4) $A \cap C \neq \Lambda$ (Algunas A son C)

(5) $C \subseteq B$ (Todas las S son B)

Obtenemos la Fig. 10, en que la parte de C que queda fuera de B ha sido sombreada horizontalmente, para mostrar que está vacía, y se han colocado cruces eslabonadas en las dos partes de la región común de A y C, para mostrar que no está vacía. El problema de interpretación se centra alrededor de la Región ①. El estudio de (4) y (5) claramente arguye a favor de la estipulación de que el som-

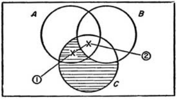


FIG. 10

breado *domina* una cruz y de aquí que la Región ① esté vacía. Estamos así en condiciones de concluir que la Región ② no está vacía, esto es, (4) y (5) implican que $A \cap (B \cap C) \neq \Lambda$ (algunas A son B y C).

Hay una constelación de circunstancias en que no deseamos decir simplemente que el sombreado domina una cruz. Cuando *cada* cruz de un eslabonamiento está cubierta por sombreado, debemos concluir que el diagrama es inconsistente y no que están completamente vacías las regiones sombreadas, pues un eslabonamiento de cruces *significa* que cuando menos una de las regiones eslabonadas no está vacía. Podemos, en realidad, usar esta circunstancia para investigar por medio de los diagramas de Venn la consistencia de una constelación de condiciones impuestas a los conjuntos. Por ejemplo, supóngase que se da de los tres conjuntos A, B y C que:

(6)	$A \subseteq C$	(Todas las A son C)
(7)	$A \cap C = \Lambda$	(Ninguna A es C)
(8)	$A \cap B \neq \Lambda$	(Algunas A son B)

Esta situación está representada por un diagrama de Venn en la Fig. 11.

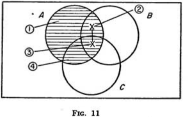


FIG. 11

El supuesto (6) nos induce a sombrear las Regiones ① y ②; el supuesto (7) nos lleva a sombrear las Regiones ③ y ④; y el supuesto (8) nos hace colocar dos cruces eslabonadas en las Regiones ② y ③. En consecuencia, los supuestos dados implican que las Regiones ② y ③ están al mismo tiempo vacías y no vacías, lo que constituye una contradicción.

Hay, por supuesto, una gran diferencia entre decir que ciertas condiciones impuestas a los conjuntos son inconsistentes y decir simplemente que implican que algún conjunto está vacío. Por ejemplo, los anteriores supuestos (6) y (7) implican que A está vacío, pero estos dos supuestos no son en sí mismos inconsistentes.

Como ya está ahora completa la notación relativa a los diagramas de Venn, es de cierto interés mostrar ahora cómo puede usarse este dispositivo para establecer la validez de los silogismos clásicos. Consideremos, a modo de ejemplo, el silogismo:

(9) Ninguna *B* es *C*
(10) Todas las *A* son *B*
(11) En consecuencia, ninguna *A* es *C*

Las premisas (9) y (10) se representan por medio de la Fig. 12.

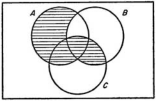


FIG. 12

Examinamos ahora el diagrama para ver si implica que ninguna *A* es *C*. Vemos en seguida que la región común a *A* y *C* está sombreada horizontalmente y concluimos que es válida la conclusión del silogismo. Todos los otros silogismos válidos pueden ser demostrados en la misma forma, pero no hay necesidad de restringir el uso de los diagramas de Venn a la prueba de validez de los argumentos que tienen la forma silogística clásica. Los diagramas de Venn pueden usarse para representar cualquier argumento que no implique más de tres conjuntos. Además, por medio de un uso cuidadoso de elipses en lugar de círculos, es posible representar gráficamente relaciones entre cuatro conjuntos, pero a menudo no es posible representar relaciones entre cinco o más conjuntos por ningún dispositivo gráfico sencillo.

EJERCICIOS

1. Sean *A* y *B* conjuntos tales que $A \cap \sim B = \Lambda$. Represéntese esta situación por un diagrama de Venn.
2. Sean *A* y *B* conjuntos tales que $A \cap \sim B = \Lambda$ y $B \cap \sim A = \Lambda$. Represéntese esta situación por un diagrama de Venn. Exprésese en forma más sencilla la relación existente entre *A* y *B*.
3. Si $A \subseteq B$ y $C \cap \sim B \neq \Lambda$, ¿qué relación existe entre *A* y *C*?
4. Trace un diagrama de Venn que represente $A \subseteq C$, y $B \cap \sim C \neq \Lambda$.
5. ¿Son las siguientes suposiciones mutuamente consistentes?

$$B \cap C = \Lambda$$
$$(A \cap C) \sim B = \Lambda$$
$$(A \cap B) \sim C = \Lambda$$
$$(A \cap B) \cup (A \cap C) \cup (B \cap C) \neq \Lambda$$

6. ¿Son las siguientes suposiciones mutuamente consistentes?

$$C \neq \Lambda$$
$$A \cap B \neq \Lambda$$
$$A \cap C = \Lambda$$
$$(A \cap B) \sim C = \Lambda$$

7. ¿Son mutuamente consistentes los siguientes supuestos?

Algunos americanos son virtuosos.
Las personas no virtuosas roban al pobre.
Algunos americanos roban al pobre.

8. Dados:

Todos los unicornios están muertos.
Ningún unicornio está muerto,

¿se podría inferir que no hay unicornios?

9. ¿Son mutuamente consistentes los supuestos del Ejercicio 8?

10. Póngase a prueba la validez de los siguientes argumentos, por medio de diagramas de Venn. Enúnciese en términos de regiones del diagrama, por qué el argumento es válido o inválido.

(a) Todos los testigos tienen prejuicios.
Algunos testigos no son mentirosos.
∴ Algunos mentirosos no tienen prejuicios.

(b) Todos los testigos tienen prejuicios.
Algunos mentirosos no tienen prejuicios.
∴ Algunos mentirosos no son testigos.

(c) Todos los mentirosos tienen prejuicios.
Algunos testigos no son mentirosos.
∴ Algunos testigos no tienen prejuicios.

(d) $A \cap B \subseteq \sim C$
$A \cup C \subseteq B$
$\therefore A \cap C = \Lambda$

(e) $A \subseteq \sim(B \cup C)$
$B \subseteq \sim(A \cup C)$
$\therefore B = \Lambda$

(f) $\sim A \subseteq \sim(B \cap C)$
$C \subseteq \sim A$
$B \subseteq A \cup \sim C$
$\therefore \sim(B \cup C) \subseteq \sim A$

§9.9. Principios elementales de las operaciones con conjuntos. Hemos visto cómo es posible usar diagramas para determinar qué conclusiones pueden obtenerse de supuestos dados. Se podría hacer una aplicación de diagramas un tanto similar, en el problema de determinar si determinadas ecuaciones o principios de otras formas, son siempre verdaderos (para todos los conjuntos). Sin embargo, un método que tiene una aplicabilidad más general y más interés desde un punto de vista matemático, es el de reducir la prueba de la verdad de una ecuación para todos los conjuntos al problema de decidir

si una fórmula correspondiente es una equivalencia tautológica, lo que siempre puede hacerse por métodos oracionales. Por ejemplo, supóngase que queremos demostrar que para todos los conjuntos A y B

(1) $$A \cup B = B \cup A.$$

Por la definición de la operación de unión sabemos que para un individuo cualquiera x

(2) $$x \in A \cup B \leftrightarrow x \in A \vee x \in B.$$

Además, es una equivalencia tautológica que

(3) $$x \in A \vee x \in B \leftrightarrow x \in B \vee x \in A.$$

Y correspondiendo a (2), tenemos:

(4) $$x \in B \cup A \leftrightarrow x \in B \vee x \in A.$$

Combinando (2) a (4), concluimos de la transitividad y simetría de la equivalencia:

(5) $$x \in A \cup B \leftrightarrow x \in B \cup A,$$

y concluimos de (5) por el principio de extensionalidad para los conjuntos, que

(6) $$A \cup B = B \cup A.$$

La inferencia de (5) a (6) es de rutina y generalmente se omite, e igualmente para la inferencia de (1) a (2). Además, estas equivalencias pueden ser presentadas en forma compacta como las identidades del Cap. 7. La prueba informal de (1) puede escribirse entonces:

$$\begin{aligned} x \in A \cup B &\leftrightarrow x \in A \vee x \in B \\ &\leftrightarrow x \in B \vee x \in A \\ &\leftrightarrow x \in B \cup A. \end{aligned}$$

Puesto que los pasos de la inferencia implican, o equivalencias tautológicas, o una de las cuatro equivalencias características:

$$\begin{aligned} x \in A \cup B &\leftrightarrow x \in A \vee x \in B \\ x \in A \cap B &\leftrightarrow x \in A \mathbin{\&} x \in B \\ x \in A \sim B &\leftrightarrow x \in A \mathbin{\&} x \notin B \\ x \in \sim A &\leftrightarrow x \notin A, \end{aligned}$$

no es necesario que cada línea lleve una razón que la justifique. Para ilustrar estos métodos con un ejemplo un poco más complicado, podemos demostrar que, para todos los conjuntos A y B

$$A \sim (A \cap B) = A \sim B.$$

DEMOSTRACIÓN:

(1) $x \in A \sim (A \cap B) \leftrightarrow x \in A \,\&\, x \notin A \cap B$
(2) $\leftrightarrow x \in A \,\&\, -(x \in A \,\&\, x \in B)$
(3) $\leftrightarrow x \in A \,\&\, (x \notin A \vee x \notin B)$
(4) $\leftrightarrow (x \in A \,\&\, x \notin A) \vee (x \in A \,\&\, x \notin B)$
(5) $\leftrightarrow x \in A \,\&\, x \notin B$
(6) $\leftrightarrow x \in A \sim B.$ Q.E.D.

Nótese que en la línea (1) eliminamos solamente una operación de conjuntos. La intersección fue a continuación eliminada en la línea (2). El reemplazo sucesivo de las operaciones de conjuntos por el conectivo oracional apropiado evita errores. Al pasar de (4) a (5) usamos la equivalencia tautológica $(\mathbf{P} \,\&\, -\mathbf{P}) \vee \mathbf{Q} \leftrightarrow \mathbf{Q}$. Para una consulta cómoda damos una lista de las identidades más útiles y de algunos otros principios estrechamente relacionados con las mismas. Por cada una de las identidades hay una correspondiente equivalencia tautológica, si tomamos a $\mathbf{P} \,\&\, -\mathbf{P}$ como correspondiente a Λ y a $\mathbf{P} \vee -\mathbf{P}$, digamos, como correspondiente a V. Se deja como ejercicio la demostración de estas identidades por los métodos oracionales que se acaban de discutir. La prueba de los principios que no son identidades es sólo ligeramente más complicada.

LISTA DE IDENTIDADES DE CONJUNTOS Y DE OTROS PRINCIPIOS

(1) $A \cup \Lambda = A$
(2) $A \cap V = A$
(3) $A \cup B = B \cup A$
(4) $A \cap B = B \cap A$
(5) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(6) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(7) $A \cup \sim A = V$
(8) $A \cap \sim A = \Lambda$
(9) $A \cup A = A$
(10) $A \cap A = A$
(11) $A \cup V = V$
(12) $A \cap \Lambda = \Lambda$
(13) $\Lambda \neq V$
(14) $\sim\sim A = A$
(15) $A = \sim B \rightarrow B = \sim A$
(16) $A \cup B \neq \Lambda \rightarrow A \neq \Lambda \vee B \neq \Lambda$
(17) $A \cap B \neq \Lambda \rightarrow A \neq \Lambda$
(18) $A \cup (B \cup C) = (A \cup B) \cup C$
(19) $A \cap (B \cap C) = (A \cap B) \cap C$
(20) $A \cup (A \cap B) = A$
(21) $A \cap (A \cup B) = A$
(22) $\sim A \neq A$
(23) $\sim(A \cup B) = \sim A \cap \sim B$
(24) $\sim(A \cap B) = \sim A \cup \sim B$
(25) $A \sim A = \Lambda$
(26) $A \sim (A \cap B) = A \sim B$
(27) $A \cap (A \sim B) = A \sim B$
(28) $(A \sim B) \sim B = A \sim B$
(29) $(A \sim B) \sim A = \Lambda$
(30) $(A \sim B) \cup B = A \cup B$
(31) $(A \cup B) \sim B = A \sim B$

Estos treinta y un principios están deliberadamente arreglados conforme a cierto orden, pues los primeros ocho pueden tomarse como axiomas de una teoría desarrollada conforme a los principios formales de la lógica de predicados de primer orden con identidad. Con adición del axioma: $(\exists A)(\exists B)(A \neq B)$ y de la definición adecuada de diferencia de conjuntos, pueden ser deducidos como teoremas los veintitrés principios restantes.[8] Cualquier modelo de estos axiomas es llamado un *álgebra booleana* y a los axiomas mismos se les da algunas veces el nombre de axiomas del álgebra de conjuntos. Un grupo de axiomas muchos más contundentes podría basarse en el concepto de membrecía en un conjunto, que no se usa en ninguno de los principios anteriores.

La mayor parte de las Ecs. (1) a (31) son tan importantes y tan frecuentemente aducidas, que se les han dado nombres especiales, algunos de los cuales son idénticos a los de las tautologías correspondientes. Otras tienen nombres derivados del hecho de que expresan importantes propiedades estructurales de las operaciones con conjuntos, como la asociatividad o la conmutatividad. La identidad (1) dice que el conjunto vacío es un *elemento de identidad* a mano derecha con respecto de la operación de unión y el (2) dice que el dominio V de individuos es un elemento de identidad a mano derecha con respecto de la intersección. (3) y (4) son las leyes *conmutativas* de la unión y la intersección. Las identidades (5) y (6) son leyes *distributivas* duales. Nótese que si comparamos la unión con la adición aritmética y la intersección con la multiplicación aritmética, el análogo de (5) no es válido en aritmética. La identidad (7) es la ley del *tercero excluido,* formulada en función de la teoría de los conjuntos; todo en un dominio dado está dentro de un conjunto o del complemento de ese conjunto. En (8) tenemos la ley de la *contradicción*; nada puede ser simultáneamente miembro de un conjunto y de su complemento. Las ecuaciones (9) y (10) expresan lo que usualmente se llama la *idempotencia* o *potencia idéntica* de la unión y la intersección. Se deja como ejercicio el problema de encontrar una operación aritmética que es idempotente. El principio (14) es la ley de la *doble negación* y (15) expresa un principio de *contraposición*. En (18) y (19) tenemos leyes *asociativas* que justifican la omisión de paréntesis, y en (20) y (21) leyes de *absorción*. Las ecuaciones (23) y (24) son las *leyes de De Morgan*.

[8] Este conjunto de axiomas se debe esencialmente a E. V. Huntington, "Sets of Independent Postulates for the Algebra of Logic", *Transactions of the American Mathematical Society*, Vol. 5 (1904), Págs 288 a 309.

Debe hacerse notar que para demostrar los principios que se refieren al conjunto vacío Λ o al dominio V de individuos, es necesario dejar sentado que para toda x

$$x \notin \Lambda,$$
$$x \in V.$$

Sea el principio (13): $\Lambda \neq V$.

Demostración: Supóngase que $\Lambda = V$. Entonces, según el principio de extensionalidad,

$$x \in \Lambda \leftrightarrow x \in V;$$

pero

$$x \notin \Lambda,$$

de donde

$$x \notin V,$$

lo cual es absurdo. Q.E.D.

Considerando un miembro arbitrario de un conjunto y usando implicaciones tautológicas en vez de equivalencias tautológicas, se establecen fácilmente principios útiles de inclusión para los conjuntos. Veamos como ejemplo:

$$A \cap B \subseteq A.$$

Demostración.

(1) $x \in A \cap B \leftrightarrow x \in A \,\&\, x \in B$

(2) $\rightarrow x \in A.$ Q.E.D.

La línea (1) es una equivalencia; el segundo miembro de (1) implica tautológicamente el segundo miembro de (2), y no es tautológicamente equivalente al mismo. La equivalencia característica usada para la inclusión es, por supuesto:

$$A \subseteq B \leftrightarrow (x)(x \in A \rightarrow x \in B).$$

Al formular demostraciones de relaciones de inclusión entre conjuntos, es necesario hacer notar un particular. Considérese:

$$(A \cap B) \sim (B \cap C) \subseteq A.$$

Demostración.

(1) $x \in (A \cap B) \sim (B \cap C) \leftrightarrow x \in A \cap B \,\&\, x \notin B \cap C$
(2) $\rightarrow x \in A \cap B$
(3) $\rightarrow x \in A \,\&\, x \in B$
(4) $\rightarrow x \in A$ Q.E.D.

El particular es que aunque la línea (3) es tautológicamente equivalente a la línea (2), usamos un signo de implicación en la línea (3) porque este signo de implicación indica la relación de (3) al lado izquierdo de (1), no al lado derecho de (2).

Las relaciones de inclusión correspondientes a todas las implicaciones tautológicas útiles que se dan en el Cap. 2 se verifican fácilmente por el método que se acaba de indicar.

EJERCICIOS

1. Demuéstrese por los métodos de este apartado la validez de cada uno de los principios (1) a (31).

2. Si la unión es interpretada como adición, la intersección como multiplicación, la diferencia de conjuntos como sustracción, el conjunto vacío como cero y el dominio V de individuos como uno, ¿cuáles de los principios (1) a (31) expresan verdades de aritmética?

3. Dése un ejemplo de operación aritmética que sea idempotente.

4. Dése un contraejemplo para demostrar que la operación de diferencia (de conjuntos) no es, en general, conmutativa.

5. Dése un contraejemplo para demostrar que la operación de diferencia no es, en general, asociativa.

6. Dése un contraejemplo para demostrar que la operación de diferencia no es, en general, distributiva con respecto de la unión, esto es, que no es verdad para todos los conjuntos A, B, C, que

$$A \sim (B \cup C) = (A \sim B) \cup (A \sim C).$$

7. ¿Es distributiva con respecto de la intersección la operación de diferencia?

8. Demuéstrese por los métodos de esta sección la validez de las siguientes relaciones de inclusión para los conjuntos cualesquiera A, B y C:

(a) $A \subseteq A \cup B$
(b) $A \sim B \subseteq A$
(c) $A \sim (B \cup C) \subseteq A \sim (B \cap C)$
(d) $(A \cap B) \sim C \subseteq (A \cup B) \sim C$

9. Usando los principios (1) a (8) como axiomas de una teoría formalizada, que usa los símbolos primitivos de operación '$\cup$', '$\cap$' y '$\sim$' y las constantes individuales 'Λ' y 'V', derívense los principios (9) a (24) como teoremas, usando como axioma adicional

(8.1) $$(\exists A)(\exists B)(A \neq B).$$

10. Añádase a los axiomas de las álgebras Booleanas del ejercicio precedente, la

DEFINICIÓN 1. $A \sim B = A \cap \sim B.$

Pruébense los principios (25) a (31).

11. Continuando el desarrollo de las álgebras Booleanas, añadimos:

DEFINICIÓN 2. $A \subseteq B \leftrightarrow A \cup B = B$.

Pruébense los siguientes teoremas (numerados consecutivamente a los treinta y un principios):

(32) $A \subseteq A$
(33) $A \subseteq B \,\&\, B \subseteq A \rightarrow A = B$
(34) $A \subseteq B \,\&\, B \subseteq C \rightarrow A \subseteq C$
(35) $\Lambda \subseteq A$
(36) $A \subseteq \Lambda \rightarrow A = \Lambda$

12. Continuando la teoría de las álgebras Booleanas, añadimos:

DEFINICIÓN 3. $A \subset B \leftrightarrow A \subseteq B \,\&\, A \neq B$.

Demostrar:

(37) $-(A \subset A)$
(38) $A \subset B \rightarrow -(B \subset A)$
(39) $A \subset B \,\&\, B \subset C \rightarrow A \subset C$
(40) $A \subset B \rightarrow A \subseteq B$
(41) $A \cap B \subset A \rightarrow A \neq B$
(42) $A \cap B \neq \Lambda \rightarrow A \sim B \subset A$

13. Haciendo referencia al último apartado del capítulo anterior, demuéstrese que ninguno de los cuatro símbolos primitivos de la teoría de las álgebras Booleanas presentados en los ejercicios anteriores es independiente de los otros.

CAPITULO 10

RELACIONES

§10.1. Pares ordenados. En el Cap. 9 nos ocupamos casi enteramente de conjuntos y del concepto de membrecía en los mismos. Los otros conceptos (como el concepto de conjunto vacío, la intersección de dos conjuntos, etc.) se *definieron* en función de estos dos (por ejemplo, el conjunto vacío se definió como un conjunto Λ, tal que para toda $x, x \notin \Lambda$; la intersección de dos conjuntos, como el conjunto de todas las cosas que pertenecen simultáneamente a los dos conjuntos dados, etc.). En cambio, en este capítulo introducimos un nuevo concepto, el de *par ordenado*, que no definimos aquí en función de los conceptos de conjunto y de membrecía.[1] Intuitivamente, un par ordenado son simplemente dos objetos dados en un orden determinado. Usamos paréntesis angulares para denotar pares ordenados. Por ejemplo, $\langle x, y\rangle$ es el par ordenado cuyo primer miembro es x y cuyo segundo miembro es y.[2] En el § 9.2 definimos dos conjuntos como idénticos cuando tienen los mismos miembros. La exigencia de identidades, en cambio, es más estricta para los pares ordenados. Dos pares ordenados son idénticos solamente cuando el primer miembro de uno es idéntico al primer miembro del otro y el segundo miembro de uno es idéntico al segundo miembro del otro. En símbolos:

(1) $$\langle x, y\rangle = \langle u, v\rangle \leftrightarrow (x = u \;\&\; y = v).$$

Tenemos, por ejemplo:

$$\{1, 2\} = \{2, 1\},$$

pero,

$$\langle 1, 2\rangle \neq \langle 2, 1\rangle.$$

[1] Esa definición es posible.

[2] La notación (x, y) se usa también ampliamente.

Podemos definir triples ordenados y, en general, n-tuples, en función de pares ordenados. Un triple ordenado, por ejemplo, es un par ordenado cuyo primer miembro es un par ordenado, esto es

(2) $$\langle x, y, z\rangle = \langle\langle x, y\rangle, z\rangle.$$

Procediendo en esta forma, definimos los cuádruples ordenados:

$$\langle x, y, z, w\rangle = \langle\langle x, y, z\rangle, w\rangle,$$

y, en general, los n-tuples ("éntuples")

$$\langle x_1, x_2, \ldots, x_n\rangle = \langle\langle x_1, x_2, \ldots, x_{n-1}\rangle, x_n\rangle.$$

En este libro trabajaremos muy poco con triples ordenados, cuádruples ordenados, etc., pero es quizás útil indicar cómo se usan las definiciones anteriores para demostrar hechos relativos a los n-tuples. Seleccionamos la sencilla tarea de demostrar que dos triples ordenados son idénticos sólo cuando son idénticos sus miembros correspondientes. Sean $\langle x, y, z\rangle$ y $\langle u, v, w\rangle$ los dos triples arbitrariamente ordenados. Está claro de la lógica de la identidad que si $x = u$, $y = v$ y $z = w$, entonces $\langle x, y, z\rangle = \langle u, v, w\rangle$. Por otra parte, comencemos con $\langle x, y, z\rangle = \langle u, v, w\rangle$. Entonces, según (2), tenemos:

$$\langle\langle x, y\rangle, z\rangle = \langle\langle u, v\rangle, w\rangle,$$

y usando (1) obtenemos

(3) $$\langle x, y\rangle = \langle u, v\rangle$$

y

$$z = w.$$

Pero (3) y (1) dan juntos:

$$x = u \text{ y } \quad y = v.$$

Hemos, pues, demostrado que

$$\langle x, y, z\rangle = \langle u, v, w\rangle \leftrightarrow (x = u \;\&\; y = v \;\&\; z = w).$$

Es importante hacer notar que la repetición del mismo elemento no añade nada en la descripción de los conjuntos, y sí en el caso de triples ordenados. Por ejemplo,

$$\{1, 2, 2\} = \{1, 2\},$$

pero $$\langle 1, 2, 2\rangle \neq \langle 1, 2\rangle,$$

puesto que $\langle 1, 2, 2\rangle = \langle\langle 1, 2\rangle, 2\rangle$ y $\langle 1, 2\rangle \neq 1$.

El concepto de *secuencia finita* puede ser definido en función de n-tuples ordenados. S es una *secuencia finita* si y sólo si hay un entero positivo n tal que S es un n-tuple ordenado. Así, por ejemplo, (Sócrates, Platón, Demócrito, Aristóteles) es una secuencia finita de filósofos griegos. En particular, es un cuádruple ordenado.

Es a menudo útil considerar el conjunto de todos los pares ordenados que pueden formarse de dos conjuntos sometidos a un orden determinado. El *producto cartesiano* (*o de cruz*) de los dos conjuntos A y B (en símbolos: $A \times B$) es el conjunto de todos los pares ordenados $\langle x, y \rangle$, tales que $x \in A$ y $y \in B$. Por ejemplo, si

$$A = \{1, 2\}$$
$$B = \{\text{Gandhi, Nehru}\}$$

entonces

$$A \times B = \{\langle 1, \text{Gandhi}\rangle, \langle 1, \text{Nehru}\rangle, \langle 2, \text{Gandhi}\rangle, \langle 2, \text{Nehru}\rangle\}.$$

§10.2. Definición de las relaciones. En contextos ordinarios a menudo hablamos de *relaciones*, que estimamos existen *entre* dos cosas o entre varias cosas. Decimos así que Isabel II está en relación de madre a hijo con el príncipe Carlos, con lo cual queremos decir que Isabel II es la madre del príncipe Carlos. O podemos decir que entre tres líneas existe una *relación de coincidencia*, con lo cual queremos decir que se cortan todas en un punto.

Para nuestras discusiones es conveniente introducir letras que designen relaciones. Así, podemos introducir la letra 'M' para la relación de madre y escribir:

$$(\text{Isabel II})\ M\ (\text{Príncipe Carlos})$$

para indicar que Isabel II es la madre del príncipe Carlos. De modo semejante, podemos desear hablar de una relación A tal que, para toda x y y, $x\ A\ y$, si y sólo si x es un antepasado de y.[3]

Al manejar una relación que es válida para tres o más cosas, es conveniente poner la letra que representa la relación y a continuación los nombres, en el orden vigente entre las cosas en que esa relación es válida. Así, por ejemplo, podemos hablar de una relación P tal que, para toda x, y, z y w, $B(x, y, z, w)$ si y solo si x debe y pesos a z para w; de manera que por ejemplo:

$$B(\text{Juan, 5, Enrique, zapatos})$$

significa que Juan debe a Enrique cinco pesos para zapatos.

[3] En el lenguaje del Cap. 4 las letras 'M' y 'A' son predicados.

Cuando hablamos de relaciones en contextos cotidianos, insistimos en que debe haber algún modo de describir la clase de conexión que existe entre cosas que decimos que están entre sí en una relación dada. Desgraciadamente, esta idea de conexión es a menudo vaga, aun cuando puede tener un significado muy preciso en casos particulares. Recuérdese que esta misma situación se suscitó en la discusión de los conectivos oracionales. La decisión tomada en el capítulo 1 fue la de permitir una conjunción de dos oraciones cualesquiera, sin importar qué desconectadas pudieran parecer intuitivamente. De modo semejante, encontramos ahora conveniente, de nuevo a causa de las dificultades de formular un preciso concepto general de conexión o dependencia, insistir en que *cualquier* conjunto de pares ordenados es una relación binaria. Así, el conjunto

$$\{\langle \text{Aristóteles}, \Lambda\rangle, \langle 7, \text{Julio César}\rangle\}$$

es una relación, aunque nadie pretendería que tiene un significado intuitivo. En consecuencia, nuestra definición de relación es

(I) *Una relación binaria es un conjunto de pares ordenados.*

De acuerdo con esta definición, la relación de *amar* es el conjunto de pares ordenados $\langle x, y\rangle$, tales que x ama a y. La relación de *ser menos que* es el conjunto de todos los pares ordenados $\langle x, y\rangle$ de números tales que, para algún número positivo z,

$$x + z = y.$$

La extensión obvia de (I) es que una relación vigente entre tres cosas es un conjunto de triples ordenados, y una relación vigente entre n cosas es un conjunto de n-tuples ordenados.

Una relación se llama 'n-aria' si sus miembros son n-tuples. Para los casos particulares $n = 2$ y $n = 3$, se usan nombres especiales y se habla de relaciones '*binarias*' y '*ternarias*'.

Puesto que una relación es un *conjunto* de n-tuples ordenados, podemos también usar la notación "$\in$" para indicar que ciertas cosas están en una relación dada. Podemos, pues, escribir:

$$\langle \text{Juan, María}\rangle \in L,$$

en vez de

$$\text{Juan } L \text{ María},$$

para indicar que Juan ama a María. De modo semejante, podemos escribir:

$$\langle \text{Jorge, María, Isabel}\rangle \in P,$$

en vez de

$$P(\text{Jorge, María, Isabel})$$

para indicar, digamos, que Jorge y María son los padres de Isabel.

Es necesario recordar que un par ordenado no es una relación, sino el conjunto consistente en el par ordenado. Por ejemplo,

⟨Tomás de Aquino, 4⟩ no es una relación;
{⟨Tomás de Aquino, 4⟩} es una relación;
{{⟨Tomás de Aquino, 4⟩}} no es una relación.

El último ejemplo no es una relación, porque el único miembro del conjunto es él mismo un conjunto, que no es un par ordenado.

Introducimos ahora una útil terminología especial para las relaciones binarias.

Si R es una relación binaria, entonces el *dominio* de R [en símbolos: $D(R)$] es el conjunto de todas las cosas x tales que, para alguna y, $\langle x, y\rangle \in R$. Así, por ejemplo, si M es la relación consistente en todos los pares $\langle x, y\rangle$ tales que x es la madre de y, entonces el dominio M es el conjunto de todas las mujeres que no carecen de hijos. Si

R_1 = {⟨Λ, Platón⟩, ⟨Jane Austen, 101⟩
⟨la novia más joven del Tibet, Richelieu⟩},

entonces

$D(R_1)$ = {Λ, Jane Austen, la novia más joven del Tibet}.

El *contradominio* (o *dominio converso*) de una relación binaria R (en símbolos: $C(R)$) es el conjunto de todas las cosas y tales que, para alguna x, $\langle x, y\rangle \in R$. El contradominio de la relación M que se acaba de considerar es el conjunto de todas las personas, puesto que toda persona tiene una madre. Si B es la relación que consiste en dos pares x, y tales que x es el hermano de y, entonces el dominio de B es el conjunto de todos los hombres que tienen cuando menos un hermano o hermana, y el contradominio es el conjunto de todas las personas que tienen cuando menos un hermano. Tenemos para la relación R_1 definida anteriormente:

$C(R_1)$ = {Platón, 101, Richelieu}.

El campo de una relación binaria R (en símbolos: $\mathscr{F}(R)$) es la unión de su dominio y su contradominio. Así, x pertenece al campo de una relación binaria R si y solo o $\langle x, z\rangle \in R$ para alguna x o $\langle z, y\rangle \in R$ para alguna y. El campo de la relación B que se acaba de

considerar es el conjunto de todas las personas pertenecientes a familias que tienen cuando menos dos hijos, uno de los cuales cuando menos es hombre. Como otro ejemplo,

$\mathscr{F}(R_1) = \{\Lambda$, Jane Austen, la novia más joven del Tibet, Platón, 101, Richelieu$\}$.

EJERCICIOS

1. Sea

$$A_1 = \{1, 2\}$$
$$A_2 = \{\Lambda\}$$
$$R = \{\langle 1, 2\rangle, \langle 2, \Lambda\rangle\}.$$

(a) ¿Es R un subconjunto del producto cartesiano $A_1 \times A_2$?
(b) ¿Es $D(R)$ un subconjunto de A_1?
(c) ¿Es $C(R)$ un subconjunto de A_2?
(d) ¿Es $\mathscr{F}(R)$ un subconjunto de $A_1 \cup A_2$?

2. ¿Cuál es el dominio, contradominio y campo de la relación consistente en ser padre?
3. ¿Qué son el dominio, contradominio y el campo de la relación consistente en ser abuelo?
4. ¿Es el dominio de la relación de ser abuelo un subconjunto propiamente dicho del dominio de la relación de ser padre?

§10.3. Propiedades de las relaciones binarias. Pasamos a tratar ahora algunas importantes propiedades de las relaciones binarias. Como lo demostrarán algunos de nuestros ejemplos y ejercicios, estas propiedades juegan un importante papel en una gran variedad de contextos científicos.

Una relación (binaria) R es *reflexiva en el conjunto A* si para toda x de A, $x\,R\,x$ (esto es, $\langle x, x\rangle \in R$). En símbolos:

$$R \text{ es reflexiva en } A \leftrightarrow (x)\ (x \in A \rightarrow x\,R\,x).$$

La relación $\leq$, por ejemplo, es reflexiva en el conjunto de todos los números reales, puesto que para todo número x, $x \leq x$. Si

$$A_1 = \{\text{Descartes, Mersenne}\},$$
$$A_2 = \{\text{Descartes}, 5\},$$

y

$$R_2 = \{\langle\text{Descartes, Descartes}\rangle, \langle\text{Mersenne, Mersenne}\rangle, \langle 5, \Lambda\rangle\},$$

entonces R_2 es reflexiva en A_1, pero no es reflexiva en A_2, puesto que el par ordenado $\langle 5, 5\rangle$ no es un miembro de R_2. La relación de amar es probablemente reflexiva en el conjunto de toda la gente; en realidad, algunos moralistas sostienen que todos nos amamos a nosotros mismos un tanto excesivamente.

Una relación R *es irreflexiva en el conjunto* A si, para toda x en A no es verdad que $x\,R\,x$. En símbolos:

$$R \text{ es irreflexiva en } A \leftrightarrow (x)\ (x \in A \rightarrow -\ (x\,R\,x)).$$

La relación de ser madre es irreflexiva en el conjunto de las personas, pues que nadie es su propia madre. La relación $<$ es irreflexiva en el conjunto de los números reales, puesto que ningún número es menor que sí mismo. Hemos visto ya que R_2 no es reflexiva en A_2, pero tampoco es irreflexiva en A_2, porque $\langle$Descartes, Descartes$\rangle \in R_2$. Considérese

$$A_3 = \{5, \text{Isabel I}\}.$$

Está claro que R_2 es irreflexiva en A_3. Del ejemplo de R_2 y A_2, debería resultar obvio que para cualquier relación R y conjunto A hay tres posibilidades

(1) R es reflexiva en A.
(2) R es irreflexiva en A.
(3) Ni (1) ni (2)

Estas tres posibilidades son mutuamente exclusivas, con una excepción: toda relación R es tanto reflexiva como irreflexiva en Λ.

Una relación R es *simétrica en el conjunto* A si para toda x y y en A, cada vez que $x\,R\,y$, entonces $y\,R\,x$. En símbolos:

$$R \text{ es simétrica en } A \leftrightarrow (x)\ (y)\ [x \in A \ \&\ y \in A \ \&\ x\,R\,y \rightarrow y\,R\,x].$$

La relación de ser primos es simétrica, pero la de amar no lo es, hecho infortunado que numerosos novelistas han puesto de relieve. La relación de ser hermano* no es simétrica, puesto que cualquier mujer que tenga un hermano suministra un ejemplo en contrario. (Nótese que en estos tres últimos ejemplos hemos omitido una referencia explícita a un conjunto en que las relaciones son o no son simétricas. Haremos esto a menudo cuando el conjunto de que se trate sea obvio, en este caso el conjunto de todas las personas). R_2 es simétrica en todos los tres conjuntos A_1, A_2 y A_3.

Una relación R es *asimétrica en el conjunto* A si, para toda x y y, en A, cada vez que $x\,R\,y$, entonces no es verdad que $y\,R\,x$. En símbolos:

$$R \text{ es asimétrica en } A \leftrightarrow (x)\ (y)\ [x \in A \ \&\ y \in A \ \&\ x\,R\,y \rightarrow -\ (y\,R\,x).$$

* En inglés se designa con una sola palabra, "cousin", primo y prima y con dos palabras "brother" y "sister", hermano y hermana, respectivamente; por este motivo considera el autor como no simétrica a la relación de ser primo y como simétrica a la de ser hermano. Está claro que dentro de este modo de ver las cosas la relación queda establecida al ras de la superficie lingüística de las palabras y no entre contenidos conceptuales. Está por demás aclarar que, si se acepta este punto de vista, en castellano serían simétricas tanto la relación de ser primo como la de ser hermano. (N. del T.)

La relación de ser madre es asimétrica, por obvias razones biológicas. Por otra parte, la relación de amar no es simétrica ni asimétrica, hecho que explica en parte el dramático interés que la misma despierta. Si necesitamos demostrar que una relación determinada no es simétrica ni asimétrica, necesitamos dar un contraejemplo definido que muestre que no es simétrica y otro diferente que demuestre que no es asimétrica. (De modo semejante necesitamos dar dos contraejemplos distintos para demostrar que una relación no es reflexiva ni irreflexiva). Por ejemplo, la relación $\leq$ en el conjunto de números no es simétrica, puesto que se tiene $1 \leq$ en el conjunto de los números no es simétrica, puesto que se tiene $1 \leq 2$, pero no $2 \leq 1$. Por otra parte, no es asimétrica, puesto que está claro que de $3 \leq 3$ no se sigue que no es cierto que $3 \leq 3$. En el primer contraejemplo reemplazamos 'x' por '1' y 'y' por '2'. En el segundo sustituimos tanto 'x' como 'y' por '3'. Al tratar de captar el sentido exacto de estas definiciones de propiedades de relaciones, es importante recordar que el *mismo* término puede remplazar a diferentes variables, tales como 'x' y 'y'.

La relación $\leq$, que no es simétrica ni asimétrica, tiene una propiedad estrechamente relacionada, que pasamos a definir. Una relación R es *antisimétrica en el conjunto A*, si para toda $x\, y$ y en A, cada vez que $x\, R\, y$ y $y\, R\, x$, entonces $x = y$. En símbolos:

R es antisimétrica en

$$A \leftrightarrow (x)\,(y)\,[x \in A \,\&\, y \in A \,\&\, x\, R\, y \,\&\, y\, R\, x \rightarrow x = y].$$

Como ya se hizo notar, $\leq$ es un ejemplo de relación antisimétrica. La relación $\subseteq$ de inclusión es un segundo ejemplo. La relación R_2 es antisimétrica en A_1, A_2 y A_3. Por otra parte, sólo en un mundo de gente completamente egocéntrica y narcisista sería amar antisimétrico. Nótese que vacuamente toda relación antisimétrica es también asimétrica. Digo 'vacuamente' porque si una relación es asimétrica en un conjunto A, entonces no hay dos objetos x y y en A tales que $x\, R\, y$ y $y\, R\, x$; esto es, nunca es cierto que $x \in A$ y $y \in A$ y $x\, R\, y$ y $y\, R\, x$. De aquí, por una simple aplicación de tablas de verdad (una implicación es verdadera cuando el antecedente es falso) siempre es cierto que si $x \in A$ y $y \in A$ y $x\, R\, y$ y $y\, R\, x$, entonces $x = y$.

Es también posible que una relación no sea simétrica, asimétrica ni antisimétrica en un conjunto. Un ejemplo es la relación de amar ya citadas varias veces. Con tal que una relación R sea válida entre cuando menos dos (no necesariamente distintos) elementos de A, las posibilidades exhaustivas y mutuamente exclusivas en el caso de las condiciones de simetría son las siguientes:

(1) R es simétrica en A
(2) R es asimétrica en A
(3) R es antisimétrica y no es simétrica ni asimétrica en A.
(4) Ni (1), (2), ni (3).

A fin de hacer las condiciones mutuamente exclusivas, tuvimos que exigir que R sea antisimétrica, pero no simétrica, así como tampoco asimétrica. Dejamos como ejercicio la construcción de una relación que sea simultáneamente simétrica y antisimétrica en un conjunto y sea válida entre dos elementos no necesariamente distintos del conjunto.

Una relación R es *transitiva en el conjunto* A si, para toda x, y y z en A, cada vez que $x\,R\,y$, y $y\,R\,z$, entonces $x\,R\,z$. En símbolos:

R es transitiva en

$$A \leftrightarrow (x)(y)(z)[x \in A \;\&\; y \in A \;\&\; z \in A \;\&\; x\,R\,y \;\&\; y\,R\,z \rightarrow x\,R\,z].$$

Las relaciones $\leq$ y $\subseteq$ son obviamente transitivas. La relación de identidad es también transitiva. Por otra parte, la relación de ser madre no lo es, puesto que si x es madre de y y y es madre de z, no puede ser que x sea también madre de z. La relación R_2 es transitiva en A_1, A_2 y A_3 en lo que puede llamarse un sentido vacío, pues no hay dos pares ordenados en R_2 que suministren un caso de prueba, por decirlo así, de transitividad, por el hecho de tener el segundo miembro de un par ordenado (la y de $x\,R\,y$) idéntico al primer miembro de otro par ordenado (la y de $y\,Rz$) y, en consecuencia, permitan la prueba que se tenga a los dos miembros restantes en la relación dada (xRz). Consideremos el caso en que se suscita esa prueba:

$A_4 = \{2, 7, \text{Goethe}\}$
$A_5 = \{2, 7, \text{Edgar Guest}\}$
$R_3 = \{\langle 2, \text{Goethe}\rangle, \langle \text{Goethe}, 7\rangle, \langle \text{Edgar Guest}, 2\rangle, \langle 2, 7\rangle\}$.

Debe resultar obvio que R_3 es transitiva en A_4 y no en A_5. Para que R_3 sea transitiva en A_5 necesitaríamos añadir el par $\langle$Edgar Guest, 7$\rangle$ a R_3. El caso de prueba para la transitividad de R_3 en A_4 es suministrado por los pares $\langle 2, \text{Goethe}\rangle$ y $\langle \text{Goethe}, 7\rangle$. Para que R_3 sea transitiva en A_4 el par $\langle 2, 7\rangle$ también debe estar en R_3. El ejemplo de R_3 puede ser engañoso. En general, no podemos decidir si una relación es transitiva en un conjunto dado considerando un solo caso de prueba; a menudo es necesario considerar varios casos y hasta decidir en forma sistemática cuál es la situación en infinidad de casos (como para $\leq$).

Relacionado con la transitividad, se tiene el concepto menos importante de intransitividad. Una relación R es *intransitiva en el conjunto A* si para toda x, y, y z de A, cada vez que $x\,R\,y$ y $y\,R\,x$, no es entonces cierto que $x\,R\,z$. En símbolos:

R intransitiva en

$$A \leftrightarrow (x)(y)(z)[x \in A \;\&\; y \in A \;\&\; z \in A \;\&\; x\,R\,y \;\&\; y\,R\,z \rightarrow -(x\,R\,z)].$$

La relación de ser madre es un ejemplo familiar de relación intransitiva. Hay una tendencia general, particularmente en las obras sobre ciencias sociales, de confundir las relaciones no transitivas y las intransitivas. Está claro que una relación puede ser no transitiva, sin ser intransitiva.

Una relación R *está conectada en el conjunto A* si por cada x y y de A, cada vez que $x \neq y$, entonces $x\,R\,y$ o $y\,R\,x$. En símbolos:

$$R \text{ conectada en } A \leftrightarrow (x)(y)(x \in A \;\&\; y \in A \;\&\; x \neq y \rightarrow x\,R\,y \vee y\,R\,x).$$

De la definición resulta obvio que una relación está conectada con un conjunto cuando conecta dos miembros cualesquiera distintos del conjunto: esto es, dados dos miembros distintos, uno está al otro en la relación dada. Las relaciones $\leq$ y $<$ están o son ambas conectadas en el conjunto de los números. Por otra parte, la relación de ser madre no está conectada en el conjunto de las personas, puesto que dadas dos personas escogidas al azar, rara vez es cierto que uno es la madre del otro. La relación R_2, no está conectada en A_1, A_2 ni en A_3. Por ejemplo, no está conectada en A_1 porque ni $\langle$Descartes, Mersenne$\rangle \in R_2$, ni $\langle$Mersenne, Descartes$\rangle \in R_2$. En cambio, la relación R_3 está conectada en A_4, pero no lo está en A_5. Para que R_3 estuviera conectada en A_5 necesitaríamos tener, o $\langle$Edgar Guest, 7$\rangle \in R_3$, o $\langle$7, Edgar Guest$\rangle \in R_3$.

Introducimos ahora una propiedad muy similar a la conexión. Una relación R está *fuertemente conectada en el conjunto A* si para toda x y y en A, o $x\,R\,y$, o $y\,R\,x$.[4] En símbolos:

$$R \text{ está fuertemente conectada en } A \leftrightarrow (x)\,(y)(x \in A \;\&\; y \in A \rightarrow x\,R\,y \vee y\,R\,x).$$

Debe resultar claro que si R está fuertemente conectada en A, entonces R está también conectada en A. La relación $\leq$ está fuertemente conectada en el conjunto de todos los números. Por otra parte, la relación $<$ no está fuertemente conectada en el conjunto de todos los números, puesto que no es verdad que $1 < 1$, esto es, si $x = y = 1$, entonces ni $x < y$, ni $y < x$.

[4] La terminología "fuertemente conectada" no se usa con amplitud. Parece que no existe un nombre generalmente convenido de esta propiedad.

Al concluir este apartado deseamos establecer dos conceptos generales que pueden ilustrarse considerando la propiedad de conexión. Sea

$$A_6 = \{2, \text{ el autor de } \textit{Hamlet}, \text{ Francisco Bacon}\}$$
$$R_4 = \{\langle 2, \text{Francisco Bacon}\rangle, \langle 2, 2\rangle\}.$$

El problema: ¿Está R_4 conectada en A_6? La respuesta no depende de la lógica, sino de la erudición sobre Shakespeare. ¿Escribió realmente Francisco Bacon las obras teatrales atribuidas a Shakespeare? Si es así, entonces resulta idéntico con el autor de *Hamlet* y R_4 está conectada en A_6. Sin embargo, la mayor parte de los eruditos no creen que Bacon escribió las obras y la respuesta más aceptable es que R_4 no está conectada en A_6. Desde nuestro punto de vista el interés de este ejemplo estriba en que pone de manifiesto que a menudo no es cuestión de lógica o de matemáticas, sino de un hecho empírico, que una relación tenga una propiedad dada como la asimetría, la transitividad o la conexión. El particular de que debe tomarse nota es que pueden usarse diferentes nombres o descripciones al hacer referencia a un solo individuo. Si Bacon hubiera escrito las obras, entonces

$$\text{Francisco Bacon} = \text{el autor de } \textit{Hamlet}$$

y al describir A_6 nos hubiéramos estado refiriendo al hombre Bacon en dos formas diferentes. Un ejemplo final que ilustra también este último punto es el siguiente. Sea

$$A_7 = \{1, 2\}$$
$$R_4 = \{\langle 1, 1\rangle, \langle 1 + 1, 2\rangle, \langle 1, 2\rangle, \langle 1 + 1, 1\rangle\}.$$

Puesto que $2 = 1 + 1$, R_4 es simultáneamente reflexiva y simétrica en A_7.

EJERCICIOS

1. Clasifique las siguientes relaciones de acuerdo con las propiedades que tienen o no tienen (por ejemplo, reflexiva, simétrica, no antisimétrica, no transitiva, etc.)

   (a) La relación de ser un abuelo en el conjunto de todas las personas.
   (b) Sea A = el conjunto de todos los números reales, y $x R y$, si y sólo si $x < y + 1$.
   (c) Sea A = el conjunto de todos los números reales, y $x R y$, si y sólo si $x < y - 1$.
   (d) Sea $A = \{1, 2, \text{Mark Twain}\}$, y $R = \{\langle 1, 2\rangle, \langle 2, 1\rangle, \langle \text{Mark Twain}, 1\rangle\}$.
   (e) Sea $A = \{3, 5, 8\}$, y $R = \{\langle 3, 3\rangle, \langle 5, 5\rangle, \langle 8, 8\rangle\}$.

(f) Sea $A = \{\text{Madison, Pinckney, } 2\}$, y $R = \{\langle 1, 1\rangle, \langle \text{Madison, Pinckney}\rangle, \langle 1, 2\rangle\}$.
(g) La relación de ser de la misma altura en el conjunto de todas las personas.
(h) La relación de ser exactamente un año más joven en el conjunto de todas las personas.
(i) La relación de divisibilidad exacta en el conjunto de los enteros positivos.

2. Enuncie las circunstancias precisas bajo las cuales una relación es simultaneamente simétrica y asimétrica en un conjunto A.

3. Sea $A = \{1, 2, \{1\}\}$.

(a) Dé un ejemplo de relación binaria que sea reflexiva y transitiva, pero no simétrica en A.
(b) Dé un ejemplo de relación binaria que sea reflexiva y simétrica, pero no transitiva en A.
(c) Dé un ejemplo de relación binaria que sea reflexiva, pero ni simétrica ni transitiva en A.
(d) Dé un ejemplo de relación binaria que no sea reflexiva, simétrica, ni transitiva en A.

4. Sea $A = \{\text{Napoleón, Wellington}\}$. ¿Puede dar un ejemplo de relación que sea irreflexiva, simétrica y transitiva en A?

5. Dé un ejemplo de relación que sea simultaneamente simétrica y antisimétrica en un conjunto, tal que sea válida entre cuando menos dos elementos del conjunto.

6. Dé un ejemplo de relación familiar que sea simultaneamente transitiva e intransitiva.

7. Cualquier balanza de brazos iguales usada para medir masas no es perfectamente sensible. Pueden encontrarse siempre tres objetos, x, y y z, tales que x equilibre exactamente a y, y y equilibre exactamente a z, pero x no equilibre exactamente a z. El triple ordenado $\langle x, y, z\rangle$ se llama una *triada intransitiva*. Dada una balanza que sea sensible no sólo a diferencias de masa mayores de, o iguales a .001 de gramo, enuncie las masas en gramos de una posible triada intransitiva para esta balanza.

§10.4. Relaciones de equivalencia. Las relaciones binarias vigentes en ciertas combinaciones de las propiedades definidas en los apartados anteriores, son significativas y útiles. En este apartado y el siguiente, definiremos algunos de los tipos más comunes.

Una relación que es reflexiva, simétrica y transitiva en el conjunto A es una *relación de equivalencia* en A. La relación de identidad es una relación de equivalencia. Lo mismo es verdad de la relación de tener la misma masa en el conjunto de los cuerpos físicos, la de tener el mismo volumen, y la existente entre dos conjuntos, si y sólo si tienen el mismo número de elementos. La relación de paralelismo entre líneas rectas es otro ejemplo. Tanto en el lenguaje matemático como en el lenguaje común las relaciones de equivalencia

se llaman a menudo relaciones de *igualdad*. Por ejemplo, como se indicó en el § 5.1, las líneas que tienen la misma longitud se llaman iguales. Para evitar la confusión común entre equivalencia en cierto aspecto de dos objetos e identidad de los mismos, siempre usamos 'igualdad' como sinónimo de 'identidad'.

Entre las relaciones de identidad y de equivalencia existe una íntima clase de conexión que tiene aplicaciones de amplias consecuencias. Describiremos una conexión, sin desarrollar sus consecuencias. Sea R una relación de equivalencia en A, y para $x \in A$, caracterizamos como sigue el conjunto $[x]$.

$$y \in [x] \leftrightarrow (y \in A \;\&\; x\,R\,y).$$

Los corchetes que contienen a 'x' se usan para designar la clase *de equivalencia R de x en A*; en otras palabras, x es la clase de equivalencia R *generada por* x. Podemos fácilmente demostrar que cada elemento de A pertenece exactamente a una clase de equivalencia R en A. Además, tenemos el importante resultado de que

(1) $$[x] = [y] \text{ si y sólo si } x\,R\,y.$$

La significación filosófica general de (1) es que formula un principio de abstracción; los objetos equivalentes en algún respecto generan clases idénticas.

Una de las razones de la importancia de este principio de abstracción es que la aplicación del mismo puede a menudo conducir a una reducción sustancial del número de entidades que se estudian, mediante el paso de los objetos a las clases equivalentes de objetos. Por ejemplo, supóngase que estamos comparando la aparente "limpieza" o ausencia de dolo de un gran número de monedas, lanzándolas al aire quince veces cada una. El resultado completo de la prueba de cada moneda puede ser representado por un 15-tuple ordenado. 'H' se usa para indicar que el resultado de un lance fue cara y 'T' para cruz. Así, por ejemplo,

$$\langle H, T, H, \ldots, H\rangle$$

indica que de los lances primero, tercero y décimoquinto se obtuvieron caras y del segundo cruz. Puesto que el resultado de cada lance es H o T, hay 2^{15} ($= 32{,}768$) resultados posibles de la prueba, esto es 32,768 15-tuples ordenados de la clase descrita. El desarrollo de una teoría para comparar estos 32,768 resultados posibles parece un tanto complicado. Pero salta a la vista una simplificación esencial. Al decidir sobre la "neutralidad lúdica" o "limpieza" de una moneda realmente sólo estamos interesados en la razón del número de lances cuyo resultado sea cara (o cruz) al número total de lances. Por

ejemplo, si una moneda produce cara quince veces seguidas, nos sentiríamos fuertemente inclinados a concluir que no era una moneda limpia, esto es, una moneda con respecto de la cual fueran prácticamente iguales la probabilidad de cara o la de cruz. Puesto que estamos interesados solamente en la razón del número de caras al número total de lances, podemos cosechar la información relativa al orden en que han ocurrido H y T. Definimos dos de nuestros tuples ordenados como *razón-equivalentes*, si en cada uno de ellos aparece H el mismo número de veces. Así

$$\langle H, H, H, T, T, T, T, H, T, T, T, H, H, T, T\rangle$$

y

$$\langle T, T, H, T, H, H, T, T, T, T, H, H, T, H, T\rangle$$

son razón-equivalentes, puesto que en cada uno de esos 15-tuples ordenados aparecen seis H. Es evidente que hay 16 clases razón-equivalentes, correspondientes a 0, 1, 2,. . . 15 ocurrencias de cara y el pequeño número de esas clases de equivalencia contrasta impresionantemente con los 32,768 tuples ordenados. En particular, es posible desarrollar reglas para rechazar o aceptar como limpia una moneda, simplemente acordando a qué clase de equivalencia pertenece su prueba. La construcción de hecho de esas reglas está más allá del alcance de nuestro estudio, puesto que es esencialmente un problema de estadística, más que de lógica.

EJERCICIOS

1. ¿Cuáles, si las hay, de las relaciones del Ejercicio 1 del apartado precedente son relaciones de equivalencia?
2. ¿Puede dar un ejemplo de una relación familiar que sea una relación de equivalencia en el conjunto de todas las personas?
3. Sea S una relación tal que $x\,S\,y$, si y sólo si x y y han nacido en el mismo estado, y sea A el conjunto de todas las personas nacidas en algún estado de los Estados Unidos.
   (a) ¿Cuántas clases hay en A, de equivalencia S?
   (b) ¿Qué clase de equivalencia S es [Truman]?
   (c) ¿Qué clase de equivalencia S es [Eisenhower]?
4. Un fabricante de dados de alta calidad prueba la limpieza de cada dado lanzándolos 100 veces y tabulando los resultados. De este modo se tienen 6^{100} resultados posibles de una prueba dada. Descríbase una relación de equivalencia intuitivamente aceptable, que agrupe los resultados posibles en un número relativamente pequeño de clases equivalentes.

§10.5. Relaciones de ordenación. En este apartado definimos las relaciones que ordenan los conjuntos en varios sentidos débiles y

fuertes. Históricamente, fueron descubiertas primero las ordenaciones fuertes, como $\leq$ y $<$ para los números reales. Las ordenaciones más débiles discutidas adelante son generalizaciones de esas ordenaciones fuertes. Las relaciones de ordenación definidas en este apartado tienen todas importantes aplicaciones tanto en matemáticas como en las ciencias empíricas, particularmente las ciencias sociales. Algunos de los ejercicios del final de este apartado ilustran posibles aplicaciones.

En este apartado no se introducen ningunas propiedades nuevas de las relaciones. Se estará prontamente en condiciones de comprender las nociones de ordenación definidas, si explícitamente se toma nota de que cada ordenación representa simplemente cierta combinación de las propiedades fundamentales definidas en el § 10.3. Comenzamos con las relaciones de ordenación lógicamente más simples.

Una relación R es una *cuasi-ordenación del conjunto A* si y sólo si R es reflexiva y transitiva en A. La relación $\subseteq$ de inclusión entre conjuntos es un ejemplo de una cuasi-ordenación (del conjunto de todos conjuntos). La relación $\subseteq$ es una cuasi-ordenación del conjunto de todos los números. La relación de ser cuando menos igual de alto es una cuasi-ordenación del conjunto de todas las personas. Puesto que cada una de estas tres relaciones tienen propiedades distintas a las de ser reflexivas y transitivas, son, como veremos, algo más que simples cuasi-ordenaciones. Para otro ejemplo, sea

$$A_1 = \{1, 2, \text{Robin Hood}\}$$

$$R_1 = \{\langle 1, 1\rangle, \langle 2, 2\rangle, \langle \text{Robin Hood}, \text{Robin Hood}\rangle, \langle 1, 2\rangle, \langle 2, \text{Robin Hood}\rangle, \langle 2, 1\rangle, \langle 1, \text{Robin Hood}\rangle\}.$$

Entonces R_1 es una cuasi-ordenación de A_1.

Una relación R es una *ordenación parcial del conjunto A* si y sólo si R es reflexiva, antisimétrica y transitiva en A. En otras palabras, una ordenación parcial de A es una cuasi-ordenación antisimétrica de A. La relación $\subseteq$ y $\leq$ que se acaban de mencionar son ordenaciones parciales. Por otra parte, R_1 no es una ordenación parcial de A_1, puesto que no es antisimétrica en A_1: $\langle 1, 2\rangle \in R_1$ y $\langle 2, 1\rangle \in R_1$, pero $1 \neq 2$. También la relación de ser cuando menos igual de alto, no es una ordenación parcial del conjunto de personas, puesto que dos hombres distintos pueden tener la misma altura y, en consecuencia, la relación no es antisimétrica.

Una ordenación parcial R de un conjunto finito A puede ser representada por un diagrama de Hasse.[5] Pequeños círculos repre-

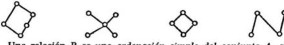
[5] Estos diagramas son llamados así por el matemático H. Hasse.

sentan elementos de A y si y puede ser alcanzada desde x, por una línea ascendente continua y no necesariamente recta,

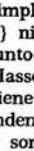
entonces $x\,R\,y$. Por ejemplo, el diagrama anterior representa el ordenamiento parcial de a, b, c, tal que $a\,R\,b$ y $c\,R\,b$. R es el conjunto

$$\{\langle a, a\rangle, \langle b, b\rangle, \langle c, c\rangle, \langle a, b\rangle, \langle c, b\rangle\}.$$

Otros ejemplos de ordenaciones parciales son dados por los otros diagramas



Una relación R es una *ordenación simple del conjunto A*, si y sólo si R es reflexiva, antisimétrica, transitiva y conectada en A. La relación $\leq$ es una ordenación del conjunto de todos los números. En otras palabras, una ordenación simple de A es una ordenación conectada parcial de A. La relación $\leq$ es una ordenación simple del conjunto de todos los números. Por otra parte, la relación $\subseteq$ de inclusión no es una ordenación simple del conjunto de todos los conjuntos, puesto que ni $\{1\} \subseteq \{2\}$ ni $\{2\} \subseteq \{1\}$; esto es, puesto que $\subseteq$ no está conectado en el conjunto de todos los conjuntos. Ninguno de los anteriores diagramas de Hasse representa una ordenación simple, puesto que cada diagrama tiene cuando menos dos elementos no conectados por una línea ascendente. De hecho, los diagramas de Hasse de ordenaciones simples son monótonos. Siempre tienen el siguiente aspecto:



esto es, un diagrama de Hasse de una ordenación simple es sólo una línea vertical.

La relación $\subset$ de inclusión propia y la relación $<$ de menos que, están conectadas estrechamente con $\subseteq$ y $\leq$, respectivamente, pero $\subset$ y $<$ no son ordenaciones parciales. Puesto que estas últimas relaciones son ejemplos importantes de relaciones de ordenación, sugieren la noción de una ordenación parcial *estricta*. Usamos la palabra 'estricta' porque la relación $<$ se llama algunas veces una desigualdad estricta. Así, una relación R es una *ordenación parcial estricta del*

conjunto A, si y sólo si R es asimétrica y transitiva en A. La relación de ser más alto que, es una ordenación parcial estricta en el conjunto de los seres humanos. Es de cierto interés notar que la relación de ser esposo es una ordenación parcial estricta en el conjunto de los seres humanos. El carácter asimétrico de la relación es obvio y la relación es transitiva en un sentido vacío, puesto que no hay tres individuos x, y y z, tales que x sea esposo de y y y esposo de z, pues esto requeriría que y fuera a la vez hombre y mujer.

Debe advertirse que en la frase 'ordenación parcial estricta' la palabra 'estricta' no funciona como adjetivo que modifique a 'ordenación parcial', pues las ordenaciones parciales estrictas no son ciertas ordenaciones parciales especiales. Excluyendo las relaciones en el conjunto vacío, ninguna ordenación puede ser a la vez una ordenación parcial y una ordenación parcial estricta, pues la primera debe ser reflexiva y la segunda irreflexiva. Por otra parte, la distinción entre ordenaciones parciales y ordenaciones parciales estrictas, es sustantivamente trivial. A cada ordenación parcial corresponde una ordenación parcial estricta única, y recíprocamente. Además, los diagramas de Hasse de una ordenación parcial y de la correspondiente ordenación parcial estricta son idénticos.

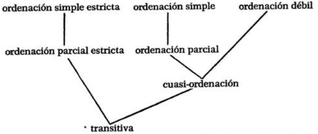
La relación entre $\leq$ y $<$ sugería ordenaciones parciales y ordenaciones parciales estrictas. También sugiere ordenaciones simples y ordenaciones simples *estrictas*. Una relación R es una *ordenación simple estricta del conjunto A*, si y sólo si R es asimétrico, transitivo y conectado en A. Como era de esperarse, una ordenación simple estricta de A es solamente una ordenación parcial estricta conectada de A. La relación $<$ en el conjunto de los números, es un ejemplo primordial de ordenación estricta simple.

Una relación R es una *ordenación débil del conjunto A*, si y sólo si R es transitivo y fuertemente conectado en A.[6] Puesto que si una relación es reflexiva y conectada, está fuertemente conectada, cualquier ordenación simple es también una ordenación débil, justamente como cualquier ordenación simple es también una ordenación parcial; pero hay ordenaciones débiles que no son ni parciales, ni simples ordenaciones; por ejemplo, la relación de ser cuando menos tan alto como, ya mencionada. También, R_1 es una ordenación débil de A_1. Si $x = y$, entonces la suposición de que $x \, R \, y$ o $y \, R \, x$ implica que $x \, R \, x$, que nos conduce a la conclusión de que cualquier ordenación débil de un conjunto es también una cuasi-ordenación de A.

El diagrama que sigue aclara las relaciones entre las diverses clases de ordenaciones introducidas en esta sección. Cuando podemos pasar de una ordenación a otra por una trayectoria continua-

[6] La expresión 'ordenación débil' no es del todo universal.

mente ascendente, esto significa que la segunda es un caso especial de la primera. Por ejemplo, toda ordenación parcial es también una cuasi-ordenación, pero, por supuesto, no toda cuasi-ordenación es una ordenación parcial. Puesto que todas las relaciones de ordenación consideradas en este apartado son transitivas, hacemos las relaciones transitivas el caso más general del diagrama.



Algunas veces, en vez de decir, por ejemplo, que la relación R es una ordenación parcial de A, es conveniente decir que A está *parcialmente ordenada* por R. Observaciones terminológicas semejantes se aplican a las otras clases de ordenaciones.

EJERCICIOS

1. En el texto no se da ningún ejemplo de una cuasi-ordenación que no sea una ordenación parcial, una ordenación débil, o una relación de equivalencia. Dése un ejemplo.

2. Determínese qué clase de relación de ordenación, si la hay, es cada una de las relaciones del Ejercicio 1 del § 10.3. Caracterícese cada relación tan plenamente como sea posible. (Por ejemplo, si una relación es una ordenación parcial, no es suficiente con decir que es una cuasi-ordenación).

3. Sea A el conjunto de matrimonios y sea $x\,P\,y$, si y sólo si el varón del matrimonio x es más bajo que el varón del matrimonio y, y la hembra del matrimonio x es más baja que la hembra del matrimonio y. ¿Es P una ordenación simple estricta de A? Si no lo es, ¿por qué no?

4. Puesto que las relaciones de equivalencias suministran una clase de ordenación, ¿dónde las pondría usted en el diagrama anterior?

5. Considérense las siguientes relaciones. Si son ordenaciones parciales u ordenaciones parciales estrictas de sus campos, trace sus diagramas de Hasse.

$R_1 = \{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 4\rangle, \langle 1, 4,\rangle\ \langle 3, 4\rangle\}$
$R_2 = \{\langle \text{Lincoln, Jefferson}\rangle, \langle \text{Jefferson, Madison}\rangle, \langle \text{Madison, Cleveland}\rangle\}$
$R_3 = \{\langle 1, 1\rangle, \langle 2, 2\rangle, \langle 3, 3\rangle, \langle 1, 2\rangle\}$
$R_4 = \{\langle 1, 2\rangle, \langle 2, 1\rangle\}$
$R_5 = \{\langle 1, 2\rangle, \langle 1, 5\rangle, \langle 5, 3\rangle, \langle 1, 3\rangle, \langle 5, 4\rangle, \langle 1, 4\rangle\}$

6. Los economistas, psicólogos, sociólogos y filósofos han estado interesados en las clases de selección que hace la gente en diversas situaciones. Han intentando establecer ciertas condiciones necesarias de racionalidad en la selección que hace la persona humana. (Véase el Ejercicio 11 del § 4.1). Si estas condiciones no se satisfacen, se sostiene que no actúa o escoge en forma intuitivamente razonable. Por ejemplo, parece razonable esperar que la relación de preferencia de un ser humano sea transitiva y asimétrica. Correspondientemente, es razonable, cuando menos como una primera aproximación, esperar que la relación de "indiferencia" de un ser humano sea una relación de equivalencia, esto es, sea reflexiva, simétrica y transitiva. Por "relación de indiferencia" queremos decir la relación existente entre dos cosas para las cuales tiene una preferencia equivalente. Así, Smith podría ser indiferente entre un Buick verde y un azul.

Resumiendo, establecemos las siguientes condiciones de racionalidad para un individuo confrontado con un conjunto A de alternativas:

(1) Su relación P de preferencia debe ser asimétrica y transitiva en A.
(2) Su relación I de inferencia debe ser una relación de equivalencia en A.
(3) Para dos alternativas en A cualesquiera, no necesariamente distintas, exactamente una de las siguientes: $x\,P\,y$, $y\,P\,x$, $x\,I\,y$.

La condición (3) simplemente dice que el individuo tiene que saber lo que desea, en el sentido de que, dadas dos alternativas cualesquiera, puede decir si prefiere una a la otra, o es indiferente entre ellas.

Nuestra definición precisa de una relación como un conjunto de pares ordenados puede ser ahora empíricamente útil, al estudiar hasta qué grado los individuos se comportan racionalmente en sus selecciones. Confrontamos a un individuo con una lista de selecciones y registramos sus respuestas, en función de pares ordenados. Aún mejor, podemos observar su comportamiento real y registrar sus decisiones.

El problema a resolver es el siguiente: dados los datos sobre cierto individuo, esto es, el registro de sus decisiones, determinar si se comporta racionalmente en el sentido anteriormente definido, y si no se comporta así, determinar qué condiciones ha violado (por ejemplo, quizás su relación I de indiferencia no es simétrica). Puesto que carece de sentido presentar un tema con una decisión entre uno y el mismo objeto, se omiten los pares de la forma $\langle x, x\rangle$, y simplemente se supone en los problemas siguientes que es reflexiva la relación I de cada tema. Nótese que corresponde al observador o al experimentador ver que cuando menos ocurra un par ordenado en la relación P o en la I, que compare cada objeto de A con cada uno de los otros.

(a) Pedimos a Jones que escoja por pares entre cuatro automóviles: Buick, Chrysler, Mercury y Studebaker. Estos automóviles constituyen el conjunto A. De sus respuestas obtenemos las siguientes relaciones P o I. ¿Son racionales esas respuestas? Si no lo son, ¿qué condiciones han violado?

$$P = \{\langle \text{Chrysler, Buick}\rangle, \langle \text{Buick, Mercury}\rangle, \langle \text{Mercury, Studebaker}\rangle, \langle \text{Buick, Studebaker}\rangle, \langle \text{Chrysler, Studebaker}\rangle\}$$
$$I = \{\langle \text{Mercury, Chrysler}\rangle, \langle \text{Chrysler, Mercury}\rangle, \langle \text{Studebaker, Chrysler}\rangle\}.$$

(b) Pedimos a los trabajadores de la fábrica que escojan entre pares de pormenores, en función de la relativa importancia que tienen para ellos.

Los pormenores son: ventilación y luz ($V \& L$), trabajo interesante (W), plan de seguro total sobre la salud (H), fondo de pensión (F), oportunidades de promoción (O).

El Sr. Smith responde como sigue:

$$P = \{\langle F, O\rangle, \langle H, O\rangle, \langle H, F\rangle, \langle H, V \& L\rangle, \langle H, W\rangle, \langle F, V \& L\rangle, \langle O, W\rangle, \langle F, W\rangle, \langle F, O\rangle\}.$$
$$I = \{\langle V \& L, W\rangle, \langle W, V \& L\rangle, \langle V \& L, O\rangle, \langle O, V \& L\rangle\}.$$

¿Procede racionalmente? Si no, ¿qué condiciones son violadas?

(c) Usando el conjunto de automóviles dado en (a), construya relaciones de preferencia racional y de indiferencia para usted mismo.

7. Los economistas usan a menudo el concepto de *espacio de artículos*. Muy simplemente, un espacio de comodidad es un conjunto de n-tuples ordenados. Para mayor brevedad de la notación podemos representado un n-tuple ordenado de mercancías o artículos para un n-tuple ordenado de números reales. Cada número representa el importe de algún artículo definido. Por ejemplo, $\langle 1, 3, 2\rangle$ puede representar un aparato de televisión, 3 piezas de pan y 2 botellas de cerveza. Esos n-tuples ordenados de mercancías se llaman a menudo *paquetes de artículos*. Ahora bien, es a menudo difícil escoger entre dos paquetes de artículos; un paquete podía contener, por ejemplo, más naranjas y menos manzanas que otro. Pero en el supuesto de que se prefiera una mayor cantidad de un artículo a una cantidad menor, puede establecerse el siguiente principio de preferencia débil:

(I) *El paquete de artículos a es preferible al paquete de artículos b, si a tiene tanto de cada artículo como b y más de cuando menos uno.*

Con base en este principio, el paquete $a_1 = \langle 1, 3, 2\rangle$ es preferible al paquete $a_2 = \langle 0, 3, 2\rangle$, puesto que a_1 tiene la misma cantidad de pan y de cerveza que a_2 y, además, un aparato de televisión.

(a) En general, ¿qué clase de relación de ordenación sobre un espacio de artículos puede derivarse del principio (I)?

(b) Sea R la relación de ordenación sobre un espacio de artículos derivada de (I). Un paquete x de un espacio de artículos S se dice que es un *punto óptimo de Pareto* si no hay y en S tal que $y R x$, esto es, tal que y es preferida a x de acuerdo con el principio (I).

Sea

$$S = \{\langle 1, 3, 2\rangle, \langle 0, 3, 2\rangle, \langle 1, 2, 2\rangle, \langle 1, 2, 5\rangle, \langle 2, 0, 4\rangle, \langle 0, 5, 1\rangle, \langle 1, 2, 0\rangle, \langle 0, 1, 0\rangle\}.$$

¿Qué elementos de S son puntos óptimos de Pareto?

(c) Sea $A = \{0, 1, 2, 3, 4, 5, 6\}$, y sea $S = A \times A \times A$. ¿Qué elementos de S son puntos óptimos de Pareto?

§10.6. Operaciones con relaciones. Si V es un dominio cualquiera de individuos, entonces *por relación universal* sobre V se quiere significar el conjunto de todos los pares ordenados $\langle x, y\rangle$ en que $x \in V$ y $y \in V$, esto es, el producto cartesiano $V \times V$.

El conjunto vacío Λ es la relación que nunca existe. Si R, por ejemplo, es la relación existente entre x y y, si y sólo si x es la madre de y, y y es la madre de x, entonces $R = \Lambda$, pues nadie es su propia madre.

Puesto que las relaciones son una clase especial de conjuntos, podemos estudiar, en su aplicación a las relaciones, las operaciones usuales con los conjuntos. Así, por ejemplo, si R y S son relaciones, entonces $R \cap S$ es la relación que consiste en la intersección de R y S: esto es, $x(R \cap S)y$, si y sólo si tanto $x\,R\,y$ como $x\,S\,y$. De modo semejante, $R \cup S$ es la unión de R y S: $x(R \cup S)y$, si y sólo si o $x\,R\,y$, o $x\,S\,y$. Y $R \sim S$ es la relación tal que $x(R \sim S)y$, si y sólo si $x\,R\,y$, y no $x\,S\,y$.

Si $x\,B\,y$, cuando x es hermano de y, y $x\,S\,y$, cuando x es hermana de y, entonces $x(B \cup S)y$, cuando x es hermano (sin consideración de sexo) de y; pero $B \cap S$ es una relación vacía. Notamos que $B \sim S = B$, pues los conjuntos de pares ordenados B y S son mutuamente exclusivos.

Si $R \subseteq S$, llamamos a R una *subrelación* de S. Así, la hermandad es una subrelación del parentesco, pues cuantas veces x es hermano de y, entonces x es un pariente de y. Toda relación es una subrelación de la relación universal sobre su propio campo.

Si hemos escogido un dominio de individuos V, y si $\cup$ es la relación universal sobre V (esto es, $\cup = V \times V$), entonces por $\sim R$, cuando R es una relación cuyo campo es un subconjunto de V, significamos simplemente $\cup \sim R$.

Además de estas operaciones, que son simplemente casos especiales de operaciones con conjuntos, es posible definir también algunas operaciones especiales con las relaciones binarias, que dependen del hecho de que son conjuntos de *pares ordenados*. Es *recíproca* de una relación R (en símbolos: R) es una relación tal que, para toda x y toda y, $x\,R\,y$, si y sólo si, $y\,R\,x$. Así pues, la recíproca de una relación se obtiene simplemente invirtiendo el orden de todos los pares ordenados que la constituyen. La recíproca de la relación H existente entre x y y si y sólo si x es el esposo de y es la relación W que existe entre y y x, si y sólo si y es la esposa de x. Así,

$$W = \breve{H}.$$

Si

$$R = \{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 4\rangle\}$$

entonces

$$\breve{R} = \{\langle 2, 1\rangle, \langle 3, 1\rangle, \langle 4, 2\rangle\}.$$

Si R y S son relaciones binarias, entonces, *producto relativo* de R y S (en símbolos: R/S) significa la relación existente entre x y y, si y sólo si existe una z tal que R es válida entre x y z, y S entre z y y. Simbólicamente,

$$x\,R/S\,y \leftrightarrow (\exists z)(x\,R\,z \,\&\, z\,S\,y).$$

Si $x\,P\,y$ significa que x es uno de los padres (la madre) de y, y $S\,y$ significa que x es hermana de y, entonces $x\,(S\,/\,P)\,y$ significa que hay una z tal que x es hermana de z y z es madre de y y, de aquí tal, que x es tía de y. La formación de productos relativos no es una operación conmutativa. Por ejemplo, con P y S como anteriormente, $P\,/\,S$ no es igual que $S\,/\,P$: la tía de un hombre rara vez es madre de su hermana.

Si $x\,P\,y$ cuando x es uno de los padres de y, entonces $x(P/P)y$, si y sólo si x es un abuelo de y; $x[(P/P)/P]y$, si y sólo si x es un bisabuelo de y; y así, sucesivamente. En la suposición de que la creación no tuvo comienzo (esto es, toda persona tuvo padres), las relaciones P, $P\,/\,P$, $(P\,/\,P)\,/\,P$, etc., son todas diferentes.

Como ejemplo final, si

$$R = \{\langle 1, 3\rangle, \langle 2, 3\rangle\},$$

$$S = \{\langle 3, 1\rangle\},$$

entonces

$$R/S = \{\langle 1, 1\rangle, \langle 2, 1\rangle\},$$

$$S/R = \{\langle 3, 3\rangle\},$$

$$(S/R)/R = \Lambda.$$

EJERCICIOS

1. Si

$x\,F\,y$	significa	x es padre de y
$x\,M\,y$	significa	x es madre de y
$x\,B\,y$	significa	x es hermano de y
$x\,S\,y$	significa	x es hermana de y

que significa decir que

(a) $x(F/M)y$
(b) $x(M/S)y$
(c) $x(M/F)y$
(d) $x(S \cup B)y$
(e) $x[S/(M \cup F)]y$
(f) $x[(B \cup S)/(M \cup F)]y$
(g) $x(\overparen{M \cup F})y$
(h) $x(\breve{M} \cup \breve{F})y$
(i) $x[(\breve{M} \cup \breve{F})/(B \cup S)]y$
(j) $x((\breve{M} \cup \breve{F})/[(B \cup S)/(M \cup F)])y$

2. Usando las letras introducidas en el Ejercicio 1, exprese lo siguiente por medio de fórmulas:

(a) x es abuelo de y
(b) x es nieto de y
(c) x es tío abuelo de y
(d) x es medio hermano de y

3. Considérense las siguientes relaciones:

$$A = \{\langle 1, 2\rangle, \langle 5, 10\rangle, \langle 2, 3\rangle, \langle 1, 3\rangle, \langle 4, 4\rangle\}$$
$$B = \{\langle 1, 1\rangle, \langle 2, 2\rangle, \langle 3, 3\rangle\}$$
$$C = \{\langle 1, 2\rangle, \langle 2, 1\rangle\}$$
$$D = \{\langle 1, 2\rangle, \langle 2, 1\rangle, \langle 1, 1\rangle, \langle 2, 2\rangle\}$$

(a) ¿Qué propiedades de reflexividad, simetría y transitividad tiene cada una de las relaciones en su propio campo?
(b) ¿Es el campo de la relación C un subconjunto del dominio de la relación A?
(c) ¿Es el dominio inverso de D un subconjunto del campo C?
(d) ¿Es cualquiera de las cuatro relaciones una relación de equivalencia en su propio campo?
(e) ¿Es B una subrelación de A?
(f) ¿Es C una subrelación de D?
(g) ¿Qué clase de relación es la relación universal sobre el campo de C?
(h) ¿Cuáles relaciones (si las hay), son idénticas a sus respectivas recíprocas?
(i) ¿Qué clase de relación es el producto relativo B/C? Liste sus miembros.
(j) Liste los miembros de C/B. ¿Es $B/C = C/B$?
(k) ¿Qué relación existe entre D/C y C/D?

4. ¿Cuál es la relación recíproca de $\leq$ para los números reales?
5. ¿Es la relación $<$ una subrelación de $\leq$?
6. ¿Qué relación de los números reales es el producto relativo $</<$? (Esto es, el producto relativo de menor que consigo mismo).
7. Sea el campo de $<$ el conjunto de todos los números reales. ¿Cuál es el campo de la relación $</<$?
8. Sea

$$A = \{\langle 1, 2\rangle, \langle 3, 4\rangle, \langle 2, 2\rangle\}$$
$$B = \{\langle 4, 2\rangle, \langle 2, 5\rangle, \langle 3, 1\rangle, \langle 1, 3\rangle\}.$$

Liste los miembros de los siguientes conjuntos:

(a) A/B
(b) $\breve{A}/\breve{B}$
(c) $A/(B \cup A)$
(d) B/A
(e) $A/(B/A)$
(f) $A \cap (A/B)$

9. Sean las relaciones A y B, definidas como ejercicios en el Ejercicio 8.

(a) ¿Es A/B un subconjunto del producto cartesiano $A \times B$?
(b) ¿Es B/A un subconjunto del producto cartesiano $D(B) \times C(B)$?

10. Sea A el conjunto de todos los números reales. Si fijamos un origen y una unidad, los puntos de un plano pueden ser considerados como el producto cartesiano $A \times A$. Una línea recta en el plano es un cierto subconjunto de $A \times A$, de donde resulta que es una relación, y podemos usar el lenguaje de este capítulo para hablar sobre operaciones con líneas.

(a) ¿Qué clase de figuras geométricas son las siguientes?
(i) la unión de dos líneas rectas;

(ii) La intersección de dos líneas rectas;
(iii) la diferencia de dos líneas rectas.

(b) Si L es una línea recta, ¿de qué tipo es su relación geométrica a su conversa $\breve{L}$?

(c) Dadas dos líneas rectas cualesquiera, ¿es su producto relativo una línea recta?

11. Una objeción a la teoría de la preferencia discutida en capítulos anteriores en diversos ejercicios es que la relación I de indiferencia no es transitiva para muchos individuos y situaciones. R. Duncan Luce ha dado un sencillo conjunto de axiomas sobre las relaciones P e I, que no requieren transitividad de I, y que pueden suministrar la base para definir dos relaciones que satisfagan los axiomas dados en el Ejercicio 6, § 10.5.[7] Sus axiomas son (para todo x, y, z, w en el conjunto de A de alternativas):

S1. $x\,I\,x$.

S2. Si $x\,P\,y$ & $y\,I\,z$ & $z\,P\,w$, entonces $x\,P\,w$.

S3. Si $x\,P\,y$ & $y\,P\,z$ & $y\,I\,w$, entonces no se tiene simultáneamente $x\,I\,w$ & $z\,I\,w$.

S4. *Exactamente una de las siguientes*: $X\ P\ y$, $y\ P\ x$, $x\ I\ y$. Definimos ahora:

(i) $x\,\overline{P}\,y \leftrightarrow x\,P\,y \vee [x\,I\,y\,\&\,(\exists z)(z \in A\,\&\,x\,I\,z\,\&\,z\,P\,y)] \vee [x\,I\,y\,\&\,(\exists z)(z \in A\,\&\,x\,P\,z\,\&\,z\,I\,y)]$

(ii) $x\,\overline{I}\,y \leftrightarrow -x\,\overline{P}\,y\,\&\,-y\,\overline{P}\,x$.

Demostración:

(a) Exactamente una de las siguientes: $x\,\overline{P}\,y$, $y\,\overline{P}\,x$, $x\,\overline{I}\,y$. (Distínganse casos en la prueba).

(b) $\overline{I}$ es reflexiva en A.

(c) $\overline{I}$ es simétrica en A.

(d) $\overline{I}$ es transitiva en A.

(e) $\overline{P}$ es transitiva en A.

[7] Duncan Luce, "Semiorders and a Theory of Utility Discrimination", Económetrica, Vol. 24 (1956), Págs. 178-191.

CAPITULO 11

FUNCIONES

§11.1. Definición. El Webster's New International Dictionary define el concepto matemático de función de esta manera:

> Una magnitud de tal modo relacionada con otra magnitud, que a los valores de esta última corresponden valores de la primera.

Las dos palabras difíciles de esta definición son 'magnitud' y 'corresponden'. ¿Qué son las magnitudes y qué significa que una cosa corresponda a otra? En el capítulo anterior sumamos el concepto de par ordenado a conjunto y membrecía, a fin de definir las relaciones. ¿Podemos evitar la suma de dos conceptos de apariencia vaga, como son los de magnitud y correspondencia, en la definición de la noción de función? Antes de decidir este asunto, veamos otra definición. Un bien conocido libro de texto[1] de cálculo diferencial e integral, afirma:

> Una variable es una cantidad a la cual pueden asignarse en una investigación un número ilimitado de valores... Cuando dos variables están relacionadas de tal modo que el valor de la primera queda determinado cuando se da el valor de la segunda, entonces se dice que la primera variable es una función de la segunda... la segunda variable... se llama la *variable independiente*, o *argumento*, y la primera variable se llama la *variable dependiente* o *función*.

Para comenzar, en esta definición la palabra 'variable' no está siendo usada como en este libro, sino más bien como sinónimo de la 'magnitud' de Webster. Esta definición no es ni mejor ni peor que la de Webster; no podemos comprender claramente ninguna de las dos definiciones, sin alguna aclaración concerniente a los elusivos conceptos de magnitud y cantidad.

[1] W. A. Granville, P. F. Smith y W. R. Longley, *Elements of the Differential and Integral Calculus*, ed. rev., Nueva York, 1941.

Una definición un tanto más refinada es la siguiente:

Una función es una regla que asigna a cada elemento de un conjunto dado un elemento único de algún otro conjunto, no necesariamente distinto.

Aquí, cuando menos, no hay expresiones imprecisas sobre magnitudes y cantidades. En muchos respectos es intuitivamente satisfactoria. Por otra parte, el concepto de regla que usa recuerda el vago concepto de dependencia o conexión que tuvimos cuidado de evitar al caracterizar los conectivos oracionales y al definir el concepto de relación.

Afortunadamente, existe un camino que elude estas dificultades. Podemos definir las funciones como relaciones binarias que relacionan cada elemento de sus dominios a un elemento *único* de sus contradominios. Más formalmente, una *función* R es una relación binaria tal que si $x\ R\ y$ y $x\ R\ z$, entonces $y = z$.

Probablemente, se está acostumbrado a pensar las funciones en términos numéricos, pero esto no es necesario. Por ejemplo, en nuestra sociedad, la relación de ser esposo es una función, puesto que una persona tiene cuando más una esposa (en un momento dado). Por otra parte, la relación de ser madre no es una función, pues una persona puede ser madre de varios hijos, pero la recíproca es una función. Será útil estudiar algunas relaciones de construcción sencillas. Sea

$$R_1 = \{\langle 1, 2\rangle, \langle \text{Madison, Pinckney}\rangle\}$$
$$R_2 = \{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle \text{Platón, Aristóteles}\rangle\}.$$

Debe resultar inmediatamente claro que R_1 es una función y R_2 no lo es, pero para evitar cualquier posible mala inteligencia, examinemos la situación con cierto detalle. El dominio de R_1 es el conjunto

$$\{1, \text{Madison}\}.$$

A cada elemento de $D(R_1)$ corresponde un elemento único del contradominio de R_1 (esto es, $C(R_1)$). Así,

$$1\ R_1\ 2 \text{ y Madison } R_1 \text{ Pinckney}$$

El dominio de R_2 es el conjunto

$$\{1, \text{Platón}\}.$$

A Platón corresponde un elemento único en $C(R_2)$, a saber, Aristóteles, pero al número uno *no* corresponde un elemento único en $C(R_2)$, pues tenemos:

$$1\ R_2\ 2 \text{ y } 1\ R_2\ 3$$

y

$$2 \neq 3.$$

Así pues, R_2 no es una función.

Hemos usado dos notaciones para las relaciones: $x\ R\ y$ y $\langle x, y\rangle \in R$. Introducimos una tercera notación usual para el caso especial de funciones. En el caso de R_1 anterior, que es una función, escribimos:

$$R_1(1) = 2,$$

$$R_1(\text{Madison}) = \text{Pinckney}.$$

Adviértase que esta notación funcional especial no puede usarse para relaciones que no sean funciones. Por ejemplo, si la aplicamos a R_2 obtenemos el siguiente absurdo:

$$R_2(1) = 2,$$

$$R_2(1) = 3,$$

y de aquí,

$$2 = 3.$$

Es a menudo conveniente usar letras minúsculas para denotar funciones. Por ejemplo,

$$f_1 = \{\langle \text{Romeo, Julieta}\rangle, \langle \text{Abelardo, Eloísa}\rangle\}$$

y

$$f_1\ (\text{Romeo}) = \text{Julieta}$$
$$f_1\ (\text{Abelardo}) = \text{Eloísa}$$

En matemáticas las funciones se definen frecuentemente por medio de ecuaciones. Por ejemplo, podría tratarse de una función f_2, tal que el dominio de f_2 sea el conjunto de todos los números reales y para cualquier número real x,

$$f_2(x) = x^2 + 2x + 5.$$

Así,

$$f_2(1) = 8 \quad \text{y} \quad f_2(2) = 13.$$

Una función no necesita ser definida por una sola ecuación. Considérese, por ejemplo, la función g, tal que el dominio de g es el conjunto de todos los números reales y para cualquier número real x

$$(1) \qquad \begin{cases} g(x) = x^2 & \text{para } x \leq 0, \\ g(x) = 5x^3 + 17 & \text{para } x > 0. \end{cases}$$

Así

$$g(-1) = 1 \quad \text{y} \quad g(1) = 22.$$

(1) es equivalente a la expresión más formal:

(2) $$g(x) = y \leftrightarrow [(x \leq 0 \rightarrow y = x^2) \,\&\, (x > 0 \rightarrow y = 5x^3 + 17)],$$

pero por razones tipográficas obvias, es costumbre usar la notación más informal, que también se escribe:

(3) $$g(x) = \begin{cases} x^2 & \text{if } x \leq 0, \\ 5x^3 + 17 & \text{if } x > 0. \end{cases}$$

Es necesario hacer hincapié en que no es formalmente suficiente con dar las ecuaciones que definen una función. Es también necesario enunciar explícitamente cuál es el dominio de la función. En la práctica matemática común este último enunciado se omite a menudo, cuando es intuitivamente claro cuál es el dominio propuesto de la función. Por ejemplo, el uso de ecuaciones definidoras como (3) por lo común trae implícitamente como consecuencia o insinúa que el dominio de la función es el conjunto de todos los números reales. No obstante, podríamos usar (3) para definir una función cuyo dominio es el conjunto de todos los números reales mayor que $=10$, digamos. La medula de la cuestión está en que, cuando el dominio de una función es un conjunto infinito, entonces es imposible listar de hecho los pares ordenados que pertenecen a la función. Puesto que la mayor parte de las funciones de interés matemático tienen dominios infinitos, la definición de una función generalmente consiste de (a) especificación de su dominio y (b) dar una regla que enuncie el valor de la función para cada elemento de su dominio.

Un segundo ejemplo de función definida para más de una ecuación es la función *h*, cuyo dominio es el conjunto de números reales positivos y que es tal, que para cualquier número positivo real *x*,

(4) $$h(x) = \begin{cases} x \text{ si } x \text{ es un entero,} \\ 1 \text{ si } x \text{ es un número racional, pero no un entero,} \\ 0 \text{ si } x \text{ no es un número racional.} \end{cases}$$

Así,

$$h(3) = 3,\ h(\tfrac{1}{2}) = 1 \quad \text{y} \quad h(\sqrt{2}) = 0.$$

Es necesario hacer también hincapié en que una función es una entidad no lingüística, sino perteneciente a la teoría de los conjuntos. Las funciones existen independientemente de cualquier lenguaje. Podemos hacer una digresión momentánea y clasificar las principa-

les entidades introducidas en este libro como lingüísticas o como pertenecientes a la teoría de los conjuntos (sinforemáticas).*

LINGUISTICO	SINFOREMATICO
Fórmulas	Individuos
Oraciones	Pares ordenados
Conectivos oracionales	Conjuntos
Variables	El conjunto vacío
Términos	Relaciones
Cuantificadores	Funciones

Las entidades lingüísticas siempre son, por supuesto, parte de algún lenguaje, mientras que las sinforemáticas no lo son por lo general. De primera intención pudiera parecer extraño decir que un individuo como Tomás de Aquino es una entidad sinforemática. La justificación se encuentra en que los individuos son ingredientes de que se construyen los conjuntos. A nivel cero tenemos los individuos; a nivel uno, conjuntos de individuos; a nivel dos, conjuntos de conjuntos de individuos, y así sucesivamente. Las unidades lingüísticas son también unidades sinforemáticas y sería más apropiado encabezar la columna de la derecha anterior, con el epígrafe "Entidades sinforemáticas que no son entidades lingüísticas".[2]

Puesto que toda función es una relación, podemos hablar del dominio y contradominio de una función. Para las funciones se usa cierta terminología especial. El dominio se llama a menudo el *dominio de definición* de la función y el contradominio se llama el *alcance de valores* de la función, para el cual usamos algunas veces la letra cursiva '$\mathcal{R}$.' Así, si f es una función,

$$C(f) = \mathcal{R}(f),$$

esto es, el contradominio de f es la misma cosa que el alcance de f. El alcance de la función h definido para (4) es el conjunto de enteros no negativos. Si f es la función cuyo dominio de definición es el conjunto de todos los números reales y que es tal para todo número real x que

$$f(x) = x^2,$$

* El adjetivo inglés "set-theoretical" sólo es posible traducirlo perifrásticamente como "perteneciente a la teoría de los conjuntos", pero esta traducción, además de ser demasiado larga, no reproduce ceñidamente el concepto inglés. El traductor se permite proponer el uso del neologismo "sinforemático", proveniente del griego το συμφόρεμα —ατος que significa montón, compuesto o conjunto. (N. del T.)

[2] El estado sinforemático de las unidades lingüísticas es un tanto complejo. Como ejemplo, considérese la variable 'x'. Por una parte, podríamos clasificar a 'x' como individuo, pero entonces, ¿qué hay que decir de las incidencias de 'x'? Ahora bien, si clasificamos las incidencias de 'x' como individuos y a la variable 'x' como el conjunto de todas esas incidencias, se suscita cierta ambigüedad con respecto del concepto de incidencia. En capítulos anteriores nos hemos referido a la incidencia de una variable en una oración, pero la ora-

entonces el alcance de f es el conjunto de todos los números reales no negativos.

Se dice algunas veces que una función traza en su alcance el *mapa* de su dominio. Así, la función h traza el mapa del conjunto de números reales en el conjunto de enteros no negativos. Usando esta clase de terminología, decimos algunas veces que f es un *mapa*, y cuando x es un elemento del dominio de f, entonces $f(x)$ es la *imagen* de f. Una función se llama ocasionalmente una *correspondencia* (lo que, incidentalmente, suministra una precisa definición de 'correspondencia'). También se llaman *transformaciones* y *operadores* ciertas funciones especiales, pero no explicaremos la base de esta denominación.

En lógica una función recibe a menudo el nombre de relación *muchos-uno*. La génesis de esta terminología debe resultar clara de la definición de función. Análogamente, una relación R es *uno-muchos* si cada vez que $y\ R\ x$ y $z\ R\ x$, entonces $y = z$. Nos referiremos a estas nociones en el apartado siguiente.

Una clase particularmente importante de funciones es la de *operaciones binarias*. Una *operación binaria con el conjunto* A es una función cuyo dominio es $A \times A$ y cuyo alcance es un subconjunto de A. Es costumbre hablar de una operación binaria *desde* $A \times A$ hasta A, y se entiende, cuando se usa este lenguaje, que el alcance de la operación necesita no ser el complemento de A, sino sólo un subconjunto de A. Por ejemplo, si N es el conjunto de enteros positivos y $+$ es la operación binaria de adición de enteros positivos, entonces decimos que $+$ es una operación binaria de $N \times N$ a N, aunque el alcance de $+$ es $N \sim \{1\}$, pues no hay dos números enteros positivos que sumen uno. Puesto que los triples ordenados son simplemente ciertos pares ordenados especiales, a saber, los pares ordenados cuyos primeros miembros son pares ordenados, podemos caracterizar las operaciones binarias como ciertas relaciones ternarias.

Una relación ternaria T es una operación binaria si y sólo si para cada x, y, z y w, si $\langle x, y, z\rangle \in T$ & $\langle x, y, w\rangle \in T$ entonces $z = w$.

EJERCICIOS

1. ¿Cuáles de las relaciones siguientes son funciones?
   (a) $\{\langle 1, 1\rangle, \langle$Judy Canova, $2\rangle, \langle$Madison, $1\rangle\}$
   (b) $\{\langle 1, 1\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle, \langle 2, 3\rangle\}$
   (c) $\{\langle 1, 1\rangle, \langle 2, 2\rangle, \langle 3, 2\rangle\}$
   (d) la relación de ser esposa
   (e) la relación de ser abuelo.

ción misma tiene muchas incidencias. El procedimiento más seguro parece ser el de comenzar con las inscripciones físicas de 'x' y construir a partir de allí. Los detalles de esta construcción no vienen al caso aquí.

2. Exprese cuál es el dominio y el alcance de cada una de las relaciones del Ejercicio 1, que sea una función.
3. ¿Cuáles de las relaciones del Ejercicio 1 del § 10.3 son funciones?
4. Sean x y y números reales. ¿Cuáles de las siguientes relaciones son funciones?

(a) La relación R, tal que $x\ R\ y$, si y sólo si $x > y$.
(b) La relación R, tal que $x\ R\ y$, si y sólo si $y \geq x + 2$.
(c) La relación R, tal que $x\ R\ y$, si y sólo si $x = y^2$.
(d) La relación R, tal que $x\ R\ y$, si y sólo si $y = x^2$.
(e) La relación R, tal que $x\ R\ y$, si y sólo si $3x + 2y = 0$.
(f) La relación R, tal que $x\ R\ y$, si y sólo si $x^2 + y^2 = 0$.
(g) La relación R, tal que $x\ R\ y$, si y sólo si $x^2 = y^2 = 0$.
(h) La relación R, tal que $x\ R\ y$, si y sólo si $x = y$.

5. Indíquese qué es el dominio y el alcance de cada una de las relaciones del Ejercicio 4 que sea una función.
6. Sean x, y y z números reales. ¿Cuál de las siguientes relaciones son operaciones binarias?[3] La relación T, tal que $\langle x, y, z\rangle \in T$, si y sólo si:

(a) $x + 2y = z$.
(b) $x < y\ \&\ y < z$.
(c) $x < y\ \&\ z = 2$.
(d) $x \cdot y = 2z$.
(e) $x + y = z - 1$.
(f) $x \cdot y \cdot z = 7$.

7. Indíquese el alcance de cada una de las relaciones ternarias del Ejercicio 6, que sea una operación binaria.

§11.2. Operaciones con las funciones. Puesto que una función es simplemente una clase especial de relación, podemos hablar del recíproco de una función. El recíproco de una función es siempre una relación, pero en general, no será una función.

Así, si

$$f = \{\langle 1, 2\rangle, \langle 2, 2\rangle\},$$

entonces

$$\breve{f} = \{\langle 2, 1\rangle, \langle 2, 2\rangle\};$$

claramente f es una función y $\breve{f}$ no lo es. Como segundo ejemplo, supóngase que f es una función tal que para cada número real x,[4]

$$f(x) = x^2;$$

si consideramos f como una relación, $\langle x, y\rangle \in f$, si y sólo si

$$x = y^2.$$

Considérese, ahora, $x = 4$. Entonces $\langle 4, 2\rangle \in \breve{f}$, y también $\langle 4, -2\rangle \in f$. De aquí que $\breve{f}$ no sea una función.

[3] Compárese el Ejercicio 2 del §8.3.
[4] Se entiende que el uso del cuantificador 'para todo número racional x' indica que el dominio de definición de f es el conjunto de todos los números reales.

Por otra parte, algunas funciones son tales que sus recíprocas son sólo funciones. Por ejemplo, si g es una función tal que para cada número real x,

$$g(x) = 2x + 4.$$

Entonces g es una función tal que

$$\breve{g}(x) = \tfrac{1}{2}x - 2.$$

Cuando una función f es tal que su relación recíproca $\breve{f}$ es también una función, entonces vemos que $\breve{f}$ es *la inversa* de f. Para marcar esta situación especial usamos '-1' como supraíndice en lugar de la notación más general '$\breve{\ }$'. Así, si la recíproca de la función de f es una función

$$f^{-1} = \breve{f}.$$

Usamos la notación '$^{-1}$' sólo cuando $\breve{f}$ es una función. Deberá advertirse que cuando existe el inverso de f (esto es, el recíproco de f es una función), tenemos las siguientes identidades: para toda x en el dominio de f

(I) $$f^{-1}(f(x)) = x,$$

y para toda x del alcance de f

(II) $$f(f^{-1}(x)) = x.$$

Los principios (I) y (II) son importantes para despejar a f y f^{-1}, como pronto lo veremos.

Al fin del apartado anterior hablábamos de relaciones muchos-uno y uno-muchos. Una relación que es al mismo tiempo muchos-uno y uno-muchos, es *uno-uno*. También se habla de una relación uno-uno como de una función uno-uno. Considérense ahora los siguientes enunciados sobre una relación R:

(1) R es una función.
(2) R es una relación muchos-uno.
(3) El recíproco de R es una función.
(4) R es una relación uno-muchos.
(5) R y su recíproco son ambos funciones.
(6) R es una función y existe su inversa.
(7) R es una relación muchos-uno y R es una relación uno-muchos.
(8) R es una relación uno-uno.
(9) R es una función uno-uno.

Existen las siguientes equivalencias

$$(1) \leftrightarrow (2),$$

$$(3) \leftrightarrow (4),$$

$$(5) \leftrightarrow (6) \leftrightarrow (7) \leftrightarrow (8) \leftrightarrow (9).$$

Cuando las funciones se definen por medio de ecuaciones, como en algunos de los ejemplos ya discutidos, se puede tener alguna dificultad en decidir si existe la inversa de una función, y, si existe, en encontrar una ecuación que defina su inversa. Sin entrar en muchos detalles, pueden ser valiosas una o dos sugestiones prácticas. Para fijar las ideas, sea f una función tal que para todo número real x

$$(1) \qquad f(x) = 5x - 4.$$

Deseamos usar (1) para encontrar f^{-1}. Puesto que (1) es válido para todo número real x, podemos sustituir x por '$f^{-1}(x)$' (aplicación de la especificación universal) y obtener:

$$(2) \qquad f(f^{-1}(x)) = 5f^{-1}(x) - 4.$$

En virtud de (II) inferimos de (2):

$$(3) \qquad x = 5f^{-1}(x) - 4.$$

Despejamos ahora '$f^{-1}(x)$' en la ecuación (3), y obtenemos la ecuación que define a f^{-1}:

$$(4) \qquad f^{-1}(x) = \tfrac{1}{5}x + \tfrac{4}{5}.$$

La estrategia de despejar a f^{-1}, tal como se ilustra en (1)-(4), tiene tres fases: (i) sustitúyase 'x' por '$f^{-1}(x)$'; (ii) aplíquese (II); (iii) despéjese '$f^{-1}(x)$' en la ecuación resultante.

La estrategia que se acaba de bosquejar tiene un defecto: su aplicación da una función inversa, aun cuando no exista esa función inversa. Considérese, por ejemplo, la función f, definida por la ecuación

$$(5) \qquad f(x) = x^2 - 1.$$

Sustituyendo 'x' por '$f^{-1}(x)$' y aplicando (II), obtenemos:

$$(6) \qquad x = f^{-1}(x)^2 - 1.$$

Despejando a '$f^{-1}(x)$' en (6), tenemos:

$$f^{-1}(x) = \pm\sqrt{x+1}.$$

Inferimos así que

$$f^{-1}(3) = 2 \quad \text{y} \quad f^{-1}(3) = -2,$$

y de aquí

$$2 = -2,$$

lo cual es absurdo. El error está en inferir (6) de (5). La sustitución de 'x' por '$f^{-1}(x)$' es permisible solamente si '$f^{-1}(x)$' es un término propiamente dicho. Si $\breve{f}$ no es una función, la expresión '$f^{-1}(x)$' no es un término, puesto que no designa una entidad única.

Algunas veces los matemáticos usan el término 'inverso' en un sentido ligeramente más amplio, dando a entender por 'la inversa' de una función f simplemente cualquier función g cuyo dominio de definición es el alcance de valores de f, y que satisface la ecuación

$$f[g(x)] = x,$$

para toda x, en el alcance de f (esto es, en el dominio de g). En este sentido, si f es una función tal que, para todo número real x,

$$f(x) = x^2,$$

entonces una función inversa de f es la función g, cuyo dominio de definición es el conjunto de números reales no negativos y que, para todo número real no negativo x, satisface la ecuación

$$g(x) = \sqrt{x}.$$

Tenemos aquí, para toda x en el alcance de f,

$$f[g(x)] = f[\sqrt{x}] = (\sqrt{x})^2 = x;$$

debe hacerse notar que aquí la ecuación

$$g[f(x)] = x$$

no se satisface para cada x del dominio de f, pues tenemos, por ejemplo,

$$g[f(-3)] = g[(-3)^2] = g(9) = \sqrt{9} = 3 \neq -3.$$

Cuando una función no tiene una inversa única los matemáticos algunas veces escogen (un tanto arbitrariamente) una inversa particular y la llaman "inversa principal" o, algunas veces, menos exactamente, el "principal valor de la inversa". Esto es especialmente común con respecto de las funciones trigonométricas. Así, cuando se trata de la función seno (esto es, con la función f tal que, para toda x, $f(x) = \operatorname{sen} x$), es costumbre introducir como inversa principal la función g, cuyo dominio de definición es el intervalo cerrado $-1 \leq x \leq +1$, y que es tal que, para cualquier número x de este intervalo, $g(x)$ es el número y numéricamente más pequeño, tal que $\operatorname{sen} y = x$; de manera que $g(\frac{1}{2}) = \pi/6$, puesto que $\operatorname{sen} \pi/6 = \frac{1}{2}$, y

$\pi/6$ es el número más pequeño numéricamente cuyo seno es ½; y de modo semejante $g(-\frac{1}{2}) = -\pi/_6$. Es costumbre denotar esta función g por 'arcsen', o 'sen^{-1}'. Notamos que

$$\operatorname{sen}[\operatorname{arcsen} x] = x,$$

para cualquier número x del dominio de definición de arcsen; por otra parte, tenemos, por ejemplo:

$$\operatorname{arcsen}\left[\operatorname{sen}\frac{5\pi}{6}\right] = \operatorname{arcsen}[\tfrac{1}{2}] = \frac{\pi}{6} \neq \frac{5\pi}{6}.$$

El producto relativo de dos funciones juega un papel importante en muchas discusiones y se le ha dado algunas veces un nombre especial: la *composición* de las funciones. Supóngase que f y g son funciones, y que x, y y z son entidades tales que

(7) $$y = f(z)$$

y

(8) $$z = g(x);$$

entonces x está a y en la relación g/f. En otros términos, de las ecuaciones (7) y (8) vemos que

$$y = f(g(x)).$$

Así, el producto relativo o composición de dos funciones es una función h tal que (cada vez que x está en el dominio de definición de g, y $g(x)$ es el dominio de definición de f)

$$h(x) = f(g(x)).$$

Esta operación de formar la composición de dos funciones es tan extensamente usada en ciertas ramas de las matemáticas que varios símbolos especiales han sido usados para la misma; nosotros usaremos un pequeño círculo '$\circ$'. Así

$$(f \circ g)(x) = f(g(x)).$$

Introducimos esta notación especial en vez de la notación del producto relativo porque el orden de 'f' y de 'g' en '$f \circ g$' es el natural que contesta a su orden en '$f(g(x))$', mientras que el orden se invierte en el producto relativo:

$$f \circ g = g/f.$$

Está claro que la asociación de funciones es asociativa pero no conmutativa.

El problema de encontrar la inversa de una función es de hecho

un caso especial del problema de encontrar "la" función g, dadas las funciones f y $f \circ g$. Al encontrar la inversa g es f^{-1}, y

$$(f \circ g)(x) = (f \circ f^{-1})(x) = f(f^{-1}(x)) = x.$$

En general la función g no es única y generalmente el problema consiste solamente en encontrar *cuando menos una* función g tal que $f \circ g$ sea idéntica a alguna función dada. Por ejemplo, dadas las funciones h y f tales que

(9) $$h(x) = x + 1$$

(10) $$f(x) = x^2,$$

encontrar una función g tal que

$$f \circ g = h.$$

Haciendo '$g(x)$' para 'x' en (10), tenemos:

$$f(g(x)) = g(x)^2,$$

de aquí que, en virtud de (9)

$$g(x)^2 = x + 1,$$

y

$$g(x) = \pm\sqrt{x + 1}.$$

Así pues, podemos tomar

$$g(x) = \sqrt{x + 1}$$

o

$$g(x) = -\sqrt{x + 1}$$

como solución de nuestro problema. La estrategia de sustituir en (10) es típica de esta clase de problema. Una estrategia algo diferente es puesta en juego cuando se nos dan funciones g y h y deseamos encontrar una función f tal que $f \circ g = h$. Por ejemplo, dadas

(11) $$h(x) = 10x^2 + 1,$$

(12) $$g(x) = 5x,$$

encontrar una f adecuada. En este caso, despejamos primeramente a g^{-1} en (12), y obtenemos:

(13) $$g^{-1}(x) = \frac{x}{5}.$$

A continuación sustituimos 'x' por '$g^{-1}(x)$' en (11), tomando, por supuesto, $h = f \circ g$:

(14) $$(f \circ g)(g^{-1}(x)) = 10(g^{-1}(x))^2 + 1.$$

Puesto que la composición es asociativa, tenemos:

$$((f \circ g) \circ g^{-1}) = (f \circ (g \circ g^{-1}))(x),$$

y, en virtud de (II)

$$f(x) = (f \circ (g \circ g^{-1}))(x).$$

De aquí inferimos de (13) y (14)

$$f(x) = 10\left(\frac{x}{5}\right)^2 + 1$$

esto es,

$$f(x) = \tfrac{2}{5}x^2 + 1.$$

Puesto que este tipo de problema implica varios pasos, puede ser útil resumir la estrategia típica: (i) despéjese a g^{-1}. (ii) Sustitúyase 'x' por $g^{-1}(x)$' en la ecuación que define a $f \circ g$. (iii) Sustitúyase a $g^{-1}(x)$ por su expresión implícita obtenida en (i), en el segundo miebro del resultado (ii). (iv) Despéjese a f.

Una pregunta natural es, ¿cuál es la situación cuando g no tiene una inversa? En ese caso, encuéntrese una función g_1 tal que

$$g(g_1(x)) = x.$$

El procedimiento para encontrar esa función g_1, es, por supuesto, idéntico al procedimiento para encontrar g^{-1}.

Finalmente, hacemos notar que, puesto que las funciones son relaciones, y las relaciones son conjuntos, se aplican a las funciones las operaciones sinforemáticas de unión, intersección y diferencia. La unión de dos funciones no es necesariamente una función, aunque la intersección de dos funciones es una función.

EJERCICIOS

1. En el Ejercicio 1 del § 11.1, ¿cuáles de las relaciones son uno-muchos, esto es, de cuáles relaciones son las funciones recíprocas?

2. En el Ejercicio 4 del § 11.1, considérense las relaciones que son funciones. ¿De cuáles existen las inversas?

3. Encuentre las inversas, si existen, de las funciones definidas para todo número real x por las siguientes ecuaciones:

(a) $f(x) = 3x + 7$
(b) $f(x) = x^2 + 1$
(c) $f(x) = -2x - 4$
(d) $f(x) = 1$

4. Sean f, g y h funciones tales que para número real x

$$f(x) = x + 2,$$
$$g(x) = x - 2,$$
$$h(x) = 3x.$$

Determínense $f \circ g, g \circ f, f \circ h, h \circ g, f \circ f, g \circ g, (f \circ h) \circ g$.

5. Sean f y g funciones tales que cada una de ellas tenga como dominio de definición el conjunto de todos los enteros positivos, y supóngase que

$$f(n) = \begin{cases} 2n & \text{si} \quad n < 3 \\ n & \text{si} \quad n \geq 3 \end{cases}$$

$$g(n) = \begin{cases} 3n & \text{si} \quad n < 3 \\ n & \text{si} \quad n \geq 3. \end{cases}$$

Determínese el dominio de definición y el alcance de $f \circ f$, $f \circ g$, $g \circ f$ y $g \circ g$.

6. Sea f_1 una función cuyo dominio de definición es A_1 y cuyo alcance es B_1; y sea f_2 una función cuyo dominio de definición es A_2 y cuyo alcance sea B_2. ¿Cuál es el dominio de definición de $f_1 \circ f_2$?

7. ¿Cuál de las siguientes identidades es válida para cualquier función f, g y h? (Cuando indicamos la inversa de una función, se supone que existe esa inversa).

(a) $f \circ f^{-1} = f/f^{-1}$
(b) $(f \circ g)^{-1} = (g/f)^{-1}$
(c) $(f \circ g) \circ h = (g/f)/h$
(d) $f \circ (g \circ h) = (h/g)/f$
(e) $f \circ g = g \circ f$
(f) $(f \circ g) \circ h = f \circ (g \circ h)$.

8. Sea f una función tal que para todo número real x,

$$f(x) = x + 2,$$

y sea g una función tal que para todo número real x,

$$g(x) = x - 2.$$

¿Cuáles de las siguientes son funciones: $f \cup g, f \cap g, f \sim g$?

9. En cada una de las siguientes encuéntrese una función g tal que $h = f \circ g$.

(a) $h(x) = 3x - 5, f(x) = 2x$
(b) $h(x) = x^2, f(x) = x^2 - 1$
(c) $h(x) = x + 1, f(x) = x^2 + 2$
(d) $h(x) = 10x - 5, f(x) = \sqrt{x}$

10. En cada una de las expresiones siguientes, encuentre una función f, tal que $h = f \circ g$.

(a) $h(x) = 3x - 5, g(x) = 2x$
(b) $h(x) = 5x, g(x) = x - 7$
(c) $h(x) = x^2, g(x) = 1 - x$
(d) $h(x) = -x^2, g(x) = -x$
(e) $h(x) = 3 - 2x, g(x) = x^2$

§11.3. Notación lambda de Church.[5] Es extremadamente importante distinguir entre una función dada y los valores de la función. Por ejemplo, supóngase que f es una función tal que, para todo número real x,

$$f(x) = x^2 + 2.$$

Entonces f es el conjunto de todos los pares ordenados de la forma $\langle x, x^2 + 2\rangle$, para x número real. Por otra parte, para cualquier otro número real particular x, $f(x)$ es simplemente el número real $x^2 + 2$. Así, no es correcto, por ejemplo, usar la frase 'la función $f(x)$', en vez de 'la función f'. Para ampliar el tema, sea

$$f_1 = \{\langle 1, \text{Washington}\rangle, \langle 8, \text{Van Buren}\rangle, \langle 13, \text{Fillmore}\rangle\}.$$

Entonces vemos fácilmente que f_1 es una función y para toda x en $D(f_1)$, $f_1(x)$ es un presidente anterior de los Estados Unidos. Carece de sentido usar la frase 'la función $f_1(x)$', puesto que ni Washington, ni Van Buren, ni Fillmore son una función.

Al introducir varias funciones especiales como ejemplos, hemos estado tratando de hacer una tajante distinción entre una función y sus valores. Nuestro propósito de consistencia con esta expresión nos ha conducido a modos de expresión, sin embargo, que son más prolongados que los usuales en matemáticas (en muchos libros de matemáticas se desprecia la distinción). Así, en los libros comunes de matemáticas, podría encontrarse el enunciado de que 'existe la inversa de la función x^3'; en cambio, nosotros tendríamos que decir 'la inversa de la función f, tal que para todo número real x

$$f(x) = x^3,$$

existe'. Es tedioso repetir frases como 'una función f tal que'.

Naturalmente, también deseamos tener nuestro pastel y comérnoslo; nos gustaría ser exactos, sin ser tediosos. Una forma parcial de llegar a esta meta es introducir la *notación lambda* de Alonzo Church, que explicaremos a continuación. La idea básica es anteponer a una expresión numérica una expresión como "λx", para obtener una expresión que designe una función. Así, el término numérico '$x + 2$', obtenemos la expresión '$(\lambda x)(x + 2)$', que designa una función f tal que para todo número real x

$$f(x) = x + 2.$$

Decimos que hemos abstraído la función $(\lambda x)(x+2)$ de la expresión '$x + 2$', y llamamos a '(λx)' un operador de abstracción. Un operador de abstracción tiene un estado lógico muy similar al de un cuantifi-

[5] Este apartado puede omitirse sin pérdida de continuidad.

cador. En particular, el uso de operadores de abstracción suministra una nueva forma de *ligar* variables. En la expresión '$(\lambda x)\ (x + 2)$' están ligadas ambas incidencias de la variable 'x'. Fijemos un dominio V de individuos. La formulación general de la abstracción (con respecto de V) es:

> *Si* v *es cualquier variable y* Φ *es cualquier término, entonces* $(\lambda v)(\Phi)$ *designa la función cuyo valor, para* $x \in V$ *es designado por el resultado de sustituir en* Φ *a* v *por un símbolo que designe a* x.

Hemos considerado un dominio fijo de individuos para simplificar la formulación. Consecuentemente, para cualquier expresión Φ, el dominio de definición de la función designada por $(\lambda v)\ \Phi$ *es* V. La suposición que los interesará de preferencia será la de que V es el conjunto de los números reales.

Sólo usamos lambda con variables, justamente como en el caso de los cuantificadores. Carece de sentido, por ejemplo, escribir $(\lambda 3)(3 + 2)$. Por otra parte, a menudo no importa qué variable en particular se use con lambda. Así,

$$(\lambda x)(x + 2) = (\lambda y)(y + 2).$$

También se permite aplicar el operador lambda a un término que contiene más de una variable; en este caso, sin embargo, no obtenemos, en general, una función definida hasta que se asigna a las variables un valor distinto del asociado con el operador lambda. Así, $(\lambda x)(ax + b)$ es una función para cada número a y b y $(\lambda x)(x^2 + y)$ es una función para toda y. Cuando hay más de una variable en Φ debe tenerse cuidado al reemplazar por otra variable, la variable usada con lambda. Así, por ejemplo, si consideramos la expresión

$$(\lambda x)(x^2 + z),$$

(que, para todo número z, designa una función) obtenemos una expresión equivalente, si reemplazamos 'x' por 'y'.

$$(\lambda y)(y^2 + z);$$

pero si reemplazamos 'x' por 'z', obtenemos:

$$(\lambda z)(z^2 + z);$$

y $(\lambda z)(z^2 + z)$ es una función idéntica a

$$(\lambda x)(x^2 + x),$$

no con

$$(\lambda x)(x^2 + z).$$

Ordinariamente, a la Φ que se antepone 'λx' contiene la variable 'x'. No obstante, no es así necesariamente; también es permitido usar la notación para las llamadas *funciones constantes*. Así, por ejemplo, $(\lambda x)(71)$ es la función f cuyo dominio es el conjunto de todos los números reales, y tal que, para todo número real x,

$$f(x) = 71.$$

La distinción entre funciones y sus valores es particularmente propensa a ser olvidada en el caso de funciones de valor constante. Pero nótese que, mientras '71' es el nombre de un número real particular, '$(\lambda x)(71)$' es el nombre de cierto conjunto de pares ordenados: el conjunto de todos los pares ordenados $\langle u, v\rangle$ tales que u es un número real y $v = 71$.

Puesto que $(\lambda x)(x^2)$ es una función, entonces

$$(\lambda x)(x^2)(3)$$

es valor de la función para el argumento 3. Como ulteriores ejemplos:

$$[(\lambda x)(x^2 + 2)](5) = 5^2 + 2 = 27$$

$$[(\lambda x)(-x^3 - 17)](-2) = -(-2)^3 - 17 = -9$$

$$[(\lambda x)(x^2 + y)](2) = 2^2 + y = 4 + y.$$

Nótese la variable libre 'y' en el último ejemplo. Debe resultar claro ahora cómo extender la notación lambda para extraer una función del término '$(\lambda x)(x^2 + y)$'. Simplemente añadimos '(λy)':

(1) $$(\lambda y)(\lambda x)(x^2 + y).$$

Se debe, por supuesto, tener cuidado de determinar valores de una función tal como (1) para argumentos dados. Usamos paréntesis para indicar cuál término, digamos '2', "reemplaza" a cuál variable. Por ejemplo,

(2) $$[(\lambda y)[(\lambda x)(x^2 + y)](2)](3) = [(\lambda y)(4 + y)](3) = 4 + 3 = 7.$$

Nuestro procedimiento consiste en trabajar de adentro hacia afuera. Comenzamos con la función de una variable $(\lambda x)(x^2 + y)$, y encontramos su valor para un elemento de su dominio. El resultado es designado por el término '$4 + y$' en (2). El '(λy)' abstrae de este término, para expresar una nueva función y nuestro paso final es el de evaluar esta nueva función para un argumento dado.

Es importante advertir que (1) no designa una función de dos variables en el sentido ordinario, sino que designa una función cuyo

dominio es el conjunto de todos los números reales y cuyo alcance es el conjunto de funciones $(\lambda x)(x^2 + y)$. Así,

$$(\lambda y)(\lambda x)(x^2 + y) = \{\langle 1, (\lambda x)(x^2 + 1)\rangle, \langle 2, (\lambda x)(x^2 + 2)\rangle, \ldots\},$$

mientras que, si f es la función de dos variables tales que para todos los números reales x y y

$$f(x, y) = x^2 + y,$$

entonces

$$f = \{\langle\langle 1, 1\rangle, 2\rangle, \langle\langle 1, 2\rangle, 3\rangle, \ldots\}.$$

La notación lambda se adapta fácilmente a permitir la abstracción de f, a saber, por reiteración de varias variables con una lambda:

$$f = (\lambda xy)(x^2 + y),$$

y

$$f(\sqrt{2}, 1) = (\lambda xy)(x^2 + y)(\sqrt{2}, 1) = (\sqrt{2})^2 + 1 = 3.$$

Como ejemplo final, la función g de las tres variables tales que para todos los números reales x, y, y z

$$g(x, y, z) = x^2 + xy + yz$$

pueda ser denotado por

$$(\lambda xyz)(x^2 + xy + yz),$$

y

(3) $\quad (\lambda xyz)(x^2 + xy + yz)(2, 4, 5) = 2^2 + 2\cdot 4 + 4\cdot 5 = 32.$

Nótese que los argumentos de la función se listan en el mismo orden en que se listan las correspondientes variables después de lambda. Así, en (3), para obtener el valor de la función correspondiente a los argumentos (2, 4, 5), reemplazamos 'x' por '2', 'y' por '4' y 'z' por '5' en la expresión de que se abstrae la función.

EJERCICIOS

1. Use la notación lambda para describir las siguientes funciones:
   (a) la función f, tal que para todo número real x,
   $$f(x) = x^3 + x^2 + 2;$$
   (b) la función f, tal que para todo número real x,
   $$f(x) = \sqrt{x} + 2;$$
   (c) la función f, tal que todos los números reales x y y,
   $$f(x) = 0;$$

(d) la función f tal que para todos los números reales x y y,

$$f(x, y) = x + y;$$

(e) la función f tal que para todos los números reales x y y

$$f(x, y) = x^2y + 2xy + y.$$

2. Evalúese lo siguiente:

(a) $[(\lambda x)(x^2 + x^4 + x^{10})](1)$

(b) $[(\lambda x)(x + x^3)](2x)$

(c) $[(\lambda x)(1 + x)](-1)$

(d) $[(\lambda z)(z + x)](2)$

(e) $[(\lambda y)([(\lambda x)(x^2 + y^3)](5))](3)$

(f) $[(\lambda y)([(\lambda x)(2xy)](3))](6)$

(g) $[(\lambda y)[(\lambda x)(x^2)](y^2)](2)$

(h) $[(\lambda x)[(\lambda y)(y^2)](\sqrt{x})](5)$

(i) $(\lambda xy)(x^2 - y^2)(3, 5)$

(j) $(\lambda xyz)(x^2y - y^2z)(1, 3, 5)$

CAPITULO 12

FUNDAMENTACION SINFOREMATICA DEL METODO AXIOMATICO

§12.1. Introducción. De todos los notables logros intelectuales de los antiguos griegos, quizás el más destacado es el desarrollo explícito del método axiomático de análisis. Los *Elementos* de Euclides, escritos alrededor de 300 años A.C., han sido probablemente el trabajo más influyente en la historia de la ciencia.[1] Toda persona culta conoce el nombre de Euclides y, a grandes rasgos, lo que hizo, esto es, que expuso la geometría en forma sistemática, a partir de un conjunto de axiomas y de postulados geométricos. Pero la mayoría de las personas estaría difícilmente en condiciones de hacer una descripción detallada del método de Euclides o de decir por qué es considerado importante desarrollar de este modo la geometría. Euclides comienza su Libro I con una lista de treinta y cinco definiciones, tres postulados y doce axiomas.[2] A continuación deduce las cuarenta y ocho proposiciones del Libro I. Las deducciones son, por supuesto, de índole más informal que formal. No hay en los *Elementos* ninguna crítica de lógica, ni ningún intento de inventariar reglas válidas de inferencia. En este particular Euclides está de acuerdo con la práctica matemática actual que discutimos en el Cap. 7. Pero en otros temas su método no está en completo acuerdo con los conceptos modernos del método axiomático. No ve claramente que el desarrollo *axiomático* de la geometría debe comenzar con algunas ideas que no se definen en función de otras. Confunde, en consecuencia, las cuestiones formales o axiomáticas con problemas relativos a la aplicación de la geometría. Por ejemplo, la Definición I del Libro I enuncia que un punto es lo que no tiene

[1] El método axiomático no es original de Euclides. Su contribución principal consistió en sistematizar los resultados de sus predecesores.
[2] Edición de Everyman's Library.

partes, o que carece de magnitud y, la Definición 2, afirma que una línea es una longitud sin anchura. Las nociones fundamentales, técnicas, no son aquí las de parte o anchura, con las cuales construimos los conceptos de punto y de línea de una forma precisa, sino más bien las de punto y de línea y son propiedades de estos últimos conceptos las que se enuncian en los postulados. Lo que Euclides trató de hacer en sus definiciones 1 y 2, fue explicar en forma intuitiva los conceptos de punto y de línea. Estas explicaciones intuitivas no son lógicamente necesarias para las demostraciones de los teoremas, pero son sugestivas en cuanto a ayudar a pensar sobre las propiedades formales de los puntos y las líneas y son también útiles para aprender a aplicar la geometría a los problemas empíricos.

En contraste con las deficiones 1 y 2, la Definición 11, que dice que un ángulo obtuso es un ángulo mayor que uno recto, tiene categoría técnica. Las pruebas de los teoremas que contienen la noción de ángulo obtuso dependen de manera formal de esta definición. Hacemos de este modo la siguiente importante distinción. Las nociones de punto y de línea son nociones *no definidas* o *primitivas* en la geometría de Euclides, mientras que la noción de ángulo obtuso es una noción *definida* o concepto.

Uno de los primeros pasos en la fundamentación axiomática de una materia o de una teoría, es el de inventariar las nociones primitivas de la teoría. No es claro, con absoluta claridad, cuáles son las nociones que deben considerarse como primitivas en el desarrollo euclidiano de la geometría. Sin embargo, hay un gran número de tratamientos axiomáticos modernos de la geometría euclidiana que son muy explícitos en este particular. Por ejemplo, podríamos tomar como primitivas las siguientes tres nociones: la noción de punto, la noción de interposición (el hecho de estar un punto entre otros dos de una línea), y la noción de equidistancia (que la distancia entre dos puntos sea igual a la distancia entre otros dos puntos dados). Usando entonces solamente el aparato de la lógica y la teoría de los conjuntos, podemos proceder a definir todas las demás nociones geométricas en función de estas tres nociones. Por ejemplo, la línea generada por dos puntos x y y se define como el *conjunto* de todos los puntos z situados entre x y y, o tales que y esté entre x y z, o tales que x esté entre z y y.

De hacerse hincapié en que las nociones primitivas de una teoría rara vez, si es que alguna, están determinadas en forma unívoca por el contenido intuitivo de la teoría. La geometría euclidiana, por ejemplo, puede desarrollarse en función de una amplia variedad de nociones primitivas, distintas de las tres mencionadas anterior-

mente. Por ejemplo, en la fundamentación axiomática del famoso matemático alemán Hilbert (1899), se introducen como primitivas las cinco nociones de punto, línea, plano, interposición y congruencia. Por otra parte, el matemático alemán Pieri publicó en el mismo año una fundamentación axiomática que usa solamente las nociones primitivas de punto y movimiento.

Un paso preliminar en la fijación de las nociones primitivas de una teoría consiste en aclarar cuáles otras teorías se supondrán en el desarrollo de la fundamentación. Para la mayor parte del trabajo axiomático de las matemáticas se supone tácitamente el desarrollo tipo de la lógica y de la teoría general de los conjuntos. Si no se hace esa suposición, entonces debe construirse un aparato completo desde los cimientos, esto es, la teoría debe ser construida dentro de un lenguaje completa y explícitamente formalizado. En el trabajo axiomático de las ciencias empíricas, como la física, la psicología y la economía, se acostumbra suponer no sólo la teoría lógica y la teoría general de los conjuntos, sino también las partes tipo de las matemáticas. Esto permite que conceptos como los de número se usen en la axiomatización de partes de la física y de la economía, digamos, y que, eso no obstante, no se consideren como primitivos.

Después de decidir cuáles otras teorías deben suponerse y de fijar las nociones primitivas de la teoría que se estudia, estamos en condiciones de enunciar los axiomas de la teoría. Los axiomas son, por supuesto, los enunciados básicos de la teoría, de los cuales podemos deducir los otros enunciados que consideramos verdaderos. Es deseable tener tan pocos axiomas como sea posible y también tomar como axiomas enunciados cuyo significado tiene un fuerte impacto intuitivo. Las únicas nociones a que se haga referencia en los axiomas deben ser las primitivas, las que se definan en función de las primitivas, y las que pertenezcan a teorías supuestas a priori. Sería, por ejemplo, impropio referirse a un objeto físico en particular, como el sol, en un axioma de geometría, puesto que es muy poco probable que, como requisito previo del desarrollo de la geometría se supusiera una teoría física y ni qué decir de una que tratara específicamente del sol. Un segundo particular que debe ser puesto de relieve es el de que, al deducir consecuencias lógicas de los axiomas, nada puede suponerse respecto de las nociones primitivas, excepto lo indicado en los axiomas.

En capítulos anteriores, en varios ejercicios y ejemplos, hemos considerado fundamentaciones axiomáticas de teorías, como la teoría de los conjuntos, que sólo necesitan suponer lógica de predicados de primer orden, con identidad. Esas teorías se llaman algunas veces *teorías con formalización tipo* (*estándar*). Muchas teorías ma-

temáticamente significativas son susceptibles de esa formalización tipo en una forma sencilla y natural, y en las últimas dos o tres décadas se ha establecido cierto número de interesantes resultados sobre varias teorías de esta clase. En el Cap. 7 se trató una formalización tipo de la aritmética de los números reales.

Desgraciadamente, cuando una teoría supone más que lógica de primer orden como ya utilizable en su enunciado y desarrollo, no es natural ni sencillo formalizar la teoría en lógica de primer orden. Por ejemplo, si al axiomatizar la geometría queremos definir las líneas como ciertos *conjuntos* de puntos, debemos trabajar dentro de un marco que ya incluye las ideas de la teoría de los conjuntos. Por supuesto, es teóricamente posible fundamentar en la lógica de primer orden, la geometría y las partes pertenecientes a la misma de la teoría de los conjuntos, pero esto es torpe e innecesariamente laborioso. Sobre todo, es redundante, pues al fundamentar axiomáticamente una amplia variedad de teorías, es necesario, o cuando menos muy conveniente, hacer uso de la teoría de los conjuntos; si se usa el método de formalización en lógica de primer orden, entonces cada fundamentación axiomática de este tipo debe incluir axiomas apropiados para la teoría de los conjuntos. Las teorías con estructuras más complicadas, como la teoría de la probabilidad necesitan usar no solamente ideas generales de la teoría de los conjuntos, sino también muchos resultados concernientes a los números reales y a funciones cuyos dominios o alcances son conjuntos de números reales. La formalización de esas teorías en lógica de primer orden es totalmente impráctica. El propósito del presente capítulo es el de suministrar una fundamentación general para la axiomatización de esas complicadas teorías y considerar con cierto detalle dos ejemplos sustantivos, uno relativo a la probabilidad y el otro a la mecánica de puntos materiales.

Aunque en este apartado ya se han discutido algunos de los aspectos generales de la axiomatización de la geometría euclidiana, por dos razones no se ha escogido como ejemplo sustantivo una axiomatización particular de la geometría. Primeramente, porque como se necesita un gran número de axiomas y como, para aquilatar la estructura de la teoría, se requiere una buena cantidad de desarrollo deductivo, las limitaciones de espacio argüían en contra de un ejemplo geométrico. Además, como una parte sustancial de la geometría euclidiana se fundamenta fácil y directamente en la lógica de primer orden, no suministraría una buena ilustración de los métodos axiomáticos, en cuanto se aplican a teorías que carecen de formalización natural en la lógica de primer orden. El propósito metodológico esencial del presente capítulo es el de demostrar que

se pueden lograr las mismas normas de claridad y precisión en la axiomatización de teorías complicadas, dentro de la teoría de los conjuntos, que las obtenidas axiomatizando teorías relativamente sencillas, directamente en lógica de primer orden. Es necesaria la comprensión de este punto de vista, para captar el amplio campo de aplicación de los métodos axiomáticos, en todos los dominios de las matemáticas y en las partes teóricas de las ciencias empíricas.

§12.2. Predicados sinforemáticos y axiomatizaciones de teorías. El meollo del procedimiento para axiomatizar teorías dentro de la teoría de los conjuntos puede ser descrito muy brevemente: axiomatizar una teoría es definir un predicado en función de nociones de la teoría de los conjuntos. Un predicado así definido se llama un predicado *sinforemático*. De hecho, en los tres capítulos anteriores ha sido definido cierto número de esos predicados, y, probablemente los dos más importantes son los predicados 'es una relación binaria' y 'es una función'.

En este capítulo no daremos una definición tajante de 'predicado sinforemático'. Nuestro objetivo es el de dilucidar por medio de ejemplos qué es lo que entra en la definición de esos predicados. Además, supondremos que la estructura sinforemática dentro de la cual trabajamos consiste no solamente en teoría general de conjuntos, tal como se discutió en capítulos anteriores, sino también en todo el aparato de las matemáticas clásicas, esto es, los números reales, funciones de números reales, derivadas e integrales de esas funciones,[3] y similares.

Al principio puede parecer difícil decidir qué cantidad de desarrollo matemático anterior es apropiado suponer en la axiomatización de una teoría dada. En la práctica, sin embargo, esta cuestión se resuelve fácilmente. Por ejemplo, al fundamentar axiomáticamente la física, es natural hacer uso de alguna parte de las matemáticas clásicas, pero al estudiar los fundamentos de, digamos, el sistema de números reales, los resultados de las matemáticas clásicas podrían usarse sólo con gran discreción y cuidado, si es que pudieran usarse. En el apartado siguiente, al estudiar la teoría de la probabilidad, es apropiado suponer como ya dado en la teoría de los conjuntos todo lo que se necesite sobre operaciones con conjuntos u operaciones con números reales; pero sería incorrecto hacer que nuestros desarrollos dependieran de resultados de la teoría intuitiva de la probabilidad.

Cualesquiera ambigüedades concernientes a lo que es ya una parte de la estructura general sinforemática, suponemos que pueden en principio ser completamente eliminadas por un desarrollo axio-

[3] Las derivadas y las integrales se necesitan solamente en el apartado final.

mático de la teoría de los conjuntos. Primero se desarrollan las propiedades de las operaciones generales con conjuntos. En seguida se construyen los números naturales, seguidos por los números reales y por el desarrollo sistemático de las matemáticas clásicas. Todo esto puede lograrse con aproximadamente unos mil teoremas y quinientas o seiscientas definiciones. La Definición 601, digamos, podría definir el predicado 'es un espacio de probabilidad'. Cualquier uso de las matemáticas clásicas en esta definición o en demostraciones de teoremas sobre espacios de probabilidad podría entonces justificarse explícita y completamente remitiendo a los correspondientes teoremas y definiciones precedentes.[4]

Pasamos ahora a hacer algunas observaciones específicas sobre la fundamentación axiomática de una teoría mediante la definición de un predicado sinforemático. Algunas observaciones son de poca monta, pero todas tienen el propósito de aclarar varias cuestiones que se suscitan en el trabajo axiomático. Nuestros ejemplos tratarán primordialmente de la teoría de las cuasi-ordenaciones o de los grupos. Para comenzar, podemos estudiar una definición de cuasi-ordenaciones ligeramente diferente de la dada en el Cap. 10. El carácter puramente sinforemático del predicado 'es una cuasi-ordenación', que se define, resulta inmediatamente evidente, al considerar los términos que aparecen en el definiente.

DEFINICIÓN A. *$\mathfrak{A}$ es una cuasi-ordenación si y sólo si hay un conjunto A y una relación binaria R, tales que* $\mathfrak{A} = \langle A, R\rangle$ *y*

Q1. *R es reflexiva en A.*
Q2. *R es transitiva en A.*

Nótese que no tendría caso reemplazar '$\mathfrak{A} = \langle A, R\rangle$' en el definiente de la definición *A* por '$\mathfrak{A} = \{A, R\}$', esto es, reemplazar el par ordenado, por un par no ordenado pues si *A* es un conjunto de pares ordenados, se suscitaría confusión sobre cuál conjunto, *A* o *R*, se considera que es la relación de ordenación. En el caso de ordenaciones parciales estrictas, hasta podríamos tener la situación anómala para algunos conjuntos *A* y *R* de que, dado $\{A, R\}$ no pudiéramos decidir si se consideraba a *A* como una ordenación parcial estricta de *R* o viceversa. Por ejemplo, sea

$$A_1 = \{\langle 1, 2\rangle\}$$
$$R_1 = \{\langle 2, 1\rangle\};$$

[4] Un grupo de matemáticos contemporáneos que escriben bajo el pseudónimo colectivo de "Bourbaki" están efectuando un desarrollo sistemático de la totalidad de las matemáticas.

entonces R_1 es una ordenación parcial estricta de A_1, y A_1 es una ordenación parcial estricta de R_1. Se originan dificultades más profundas para algunas teorías complicadas, cuando se usan conjuntos no ordenados, en vez de n-tuples ordenados.

La forma de la Definición A satisface la regla de la definición adecuada de los símbolos de relación dada en el Cap. 8, pero discrepa ligeramente del estilo dominante en las matemáticas modernas. Para ilustrar otras posibilidades, consideramos otras alternativas. Una primera alternativa es la de definir el predicado de dos posiciones 'es una cuasi-ordenación de', que es lo que hicimos en el Cap. 10.

DEFINICIÓN B. *R es una cuasi-ordenación de A, si y sólo si R es una ordenación binaria, reflexiva y transitiva en A.*

Al comparar las Definiciones A y B podría parecer de primera intención que todas las ventajas de la simplicidad están con B, puesto que en el definiente no se requieren cuantificadores o referencias a pares ordenados. Sin embargo, la ventaja de A es que muestra un procedimiento uniforme, aplicable a toda teoría, pues la redacción de definiciones como B se hace torpe cuando el predicado definido es de más de dos posiciones. Así, correspondiendo a B, tendríamos para sistemas de mecánica de puntos materiales, definida en el §12.5, algo como:

P es un sistema de mecánica de puntos materiales con respecto de T, s, m, f y g, si y sólo si ...

Una segunda alternativa, que se acerca mucho a la práctica matemática tipo, es usar una definición condicional.

DEFINICIÓN C. *Sea A un conjunto y R una relación binaria. Entonces* $\langle A, R\rangle$ *es una cuasi-ordenación si y sólo si*

Q1. *R es reflexiva en A.*
Q2. *R es transitiva en A.*

En el caso de C, la estructura puramente sinforemática de las cuasi-ordenaciones es la hipótesis de la definición condicional. Esas definiciones condicionales parecen muy naturales, y las usaremos en lo que sigue, pero promueven la continua comisión de cierto tipo de pequeño error, que puede ser ilustrado por C. Supóngase que alguien enuncia como teorema:

Si $\langle A, R\rangle$ es una cuasi-ordenación, entonces A es un subconjunto del campo de R.

Esta aserción parece ser una consecuencia obvia de C y de hechos familiares sobre las relaciones. Sin embargo, la dificultad está en

que, puesto que no se enuncia en la hipótesis del teorema que A es un conjunto y R una relación binaria, no podemos usar C para inferir que R es reflexiva en A, de modo que podamos obtener la conclusión deseada. En otras palabras, puesto que no se satisface la cláusula condicional de la Definición C, no podemos usar significativamente C en una inferencia. Pero en muchos contextos esta crítica es una sutileza. Es perfectamente obvio que se pretende que el teorema enunciado se aplique solamente a las entidades sinforemáticas adecuadas.

En discusiones que tratan de entidades que tienen la misma estructura general sinforemática, a menudo se resuelve el problema que se acaba de mencionar mediante el recurso de especificar al comienzo de la discusión esta estructura general. Por ejemplo, podríamos definir una *estructura simple de relación* como un par ordenado consistente en un conjunto y en una relación binaria. Entonces todas nuestras definiciones de relaciones de ordenación son para esas estructuras. Así

DEFINICIÓN D. *Una estructura simple de relación* $\langle A, R\rangle$ *es una cuasi-ordenación si y sólo si* R *es reflexiva y transitiva en* A.

Al considerar grupos y estructuras algebraicas relacionados, podemos considerar álgebras consistentes en un conjunto no vacío A y en una operación binaria $\circ$ de $A \times A$ a A, esto es, $\circ$ es una función cuyo dominio es $A \times A$ y cuyo alcance es un subconjunto de A. Definimos entonces:

DEFINICIÓN E. *Un álgebra* $\langle A, \circ\rangle$ *es un grupo si y sólo si, para cada* x, y *y* z *en* A

A1. $x \circ (y \circ z) = (x \circ y) \circ z$.

A2. Hay una w en A tal que

$$x = y \circ w.$$

A3. Hay una w en A tal que

$$x = w \circ y.$$

(Aquí hemos usado la formulación ya expuesta de los axiomas para grupos en el Ejercicio 5 del §5.2.) Las nuevas Definiciones C, D y E violan todas las reglas de las definiciones condicionales expuestas en el Cap. 8. Por ejemplo, en el caso de C, el definiendo usa el término '$\langle A, R\rangle$', en que se usaría una sola variable. Para rectificar este error y a continuación señalar cómo se subsana cómodamente en la práctica, podemos considerar la Definición E. Se puede formularla de nuevo.

DEFINICIÓN F. *Si* $\mathfrak{A} = \langle A, \circ \rangle$ *y* $\langle A, \circ \rangle$ *es una álgebra, entonces* $\mathfrak{A}$ *es un grupo, si y sólo si*. . .

(La letra gótica '$\mathfrak{A}$' se usa en deferencia a un uso común en los escritos sobre la materia. La razón para usarla es ésta: *A* denota el conjunto básico del grupo, mientras que $\mathfrak{A}$ es el conjunto básico, juntamente con la operación.) Es tedioso repetir continuamente la larga cláusula condicional de F. La abreviatura tipo consiste en escribir en su lugar:

DEFINICIÓN G. *Un álgebra* $\mathfrak{A} = \langle A, \circ \rangle$ *es un grupo, si y sólo si*. . .

Si por cualquier razón se desea una formulación exacta, puede siempre ampliarse al estilo de la Definición G la frase "Una álgebra $\mathfrak{A} = \langle A, \circ \rangle$" de la Definición F.

Con respecto de la definición de grupos debe observarse que muchos matemáticos prefieren una definición como:

DEFINICIÓN H.' *Un conjunto A es un grupo con respecto de la operación binaria* $\circ$ *si y sólo si*. . .

Pero aquí el predicado es tan largo que es costumbre referirse al conjunto *A* sólo como un grupo, lo que es literalmente falso y puede inducir a error. La Definición G tiene simultáneamente las virtudes de brevedad y explicitez.

En los contextos matemáticos ordinarios se formulan frecuentemente definiciones en forma matemática, pero esta formulación matemática no contiene ningún compromiso real matemático, esto es, compromisos para probar aserciones sobre expresiones de algún lenguaje dado y fijo.[5] La Definición G, por ejemplo, podría formularse:

DEFINICIÓN I. *Un álgebra* $\mathfrak{A} = \langle A, \circ \rangle$ *es un grupo si y sólo si se satisfacen los tres axiomas siguientes*. . .

Las pruebas matemáticas sobre grupos no se usarían en ninguna propiedad de forma explícita o profunda de la noción matemática de satisfacción que se usa en la Definición I.

Cuando se axiomatiza una teoría por definición de un predicado sinforemático, por *modelo* de la teoría se quiere decir simplemente una entidad que satisface el predicado. Para la teoría de las cuasi-ordenaciones podemos enunciar la cuestión trivialmente como sigue. Si $\langle A, R \rangle$ es una cuasi-ordenación, entonces $\langle A, R \rangle$ es un modelo

[5] Hablando en términos generales, las metamatemáticas son la rama de las matemáticas que investigan la estructura de lenguajes formalizados o teorías y su relación con otras entidades matemáticas. Muchos filósofos tienden a llamar *semántica* y *sintaxis lógica* y no metamatemáticas al estudio de los lenguajes formalizados.

para la teoría de la cuasi-ordenación. Correspondientemente, si un álgebra $\langle A, \circ \rangle$ es un grupo, entonces $\langle A, \circ \rangle$ es un modelo para la teoría de los grupos. Cuando la teoría de los grupos se axiomatiza directamente en lógica de primer orden, la noción de modelo se define de manera que las mismas entidades sinforemáticas sean modelos de la teoría así formulada y de modo semejante para otras teorías que pueden ser axiomatizadas directamente en lógica de primer grado, o definiendo un predicado sinforemático: las dos axiomatizaciones tienen las mismas entidades como modelos.

Durante las dos últimas décadas, las expresiones 'modelo' y 'modelo matemático' han sido ampliamente usadas, particularmente en las ciencias de la conducta. Estas expresiones parecen usarse cuando menos en tres sentidos diferentes. El sentido que tiene en lógica se acaba de describir. Un segundo significado de 'modelo' para los economistas matemáticos está estrechamente relacionado con el anterior: el modelo de una teoría es el conjunto de todos los modelos de la teoría, en el sentido del lógico. Lo que los lógicos llaman un modelo, se designà como una *estructura*. En esta terminología, si un álgebra $\langle A, \circ \rangle$ es un grupo, entonces es una estructura de la teoría de los grupos. El tercer significado de 'modelo' es el más popular en las ciencias empíricas y lo que se ha significado con el término 'teoría' en páginas precedentes. En este sentido, dar un modelo matemático para alguna rama de las ciencias empíricas equivale a enunciar una exacta teoría matemática. En esos contextos empíricos la palabra 'teoría' a menudo se reserva para enunciados no matemáticos, relativamente inexactos, sobre las ideas fundamentales de un dominio dado de la ciencia. La diferencia importante entre los primeros dos sentidos de modelo y el tercero estriba en que solamente en el tercer sentido son modelos las entidades lingüísticas.

Cuando se formalizan teorías en lógicas de primer orden, los teoremas que relacionan diversos modelos de la teoría son necesariamente metamatemáticas en su enunciado y su demostración.[6] En contraste, pueden enunciarse en directa forma matemática teoremas que comparan diversos modelos de una teoría, cuando ésta se axiomatiza definiendo un predicado sinforemático. Considérese, por ejemplo, el siguiente teorema sobre cuasi-ordenaciones.

TEOREMA 1. *Si* $\langle A, R_1 \rangle$ *y* $\langle A, R_2 \rangle$ *son cuasi-ordenaciones, entonces* $\langle A, R_1 \cap R_2 \rangle$ *es una cuasi-ordenación·*

[6] Hablando en términos generales los métodos metamatemáticos no son usados por los matemáticos cuando pueden evitarlo, pues su aplicación exacta requiere la tarea difícil y a menudo tediosa de trabajar con un lenguaje completamente especificado y formalizado. Esta observación no tiene el propósito de desvalorizar en modo alguno la importancia de las metamatemáticas. Muchos resultados importantes pueden establecerse sólo por medios metamatemáticos.

DEMOSTRACIÓN. Necesitamos demostrar que $R_1 \cap R_2$ es reflexiva y transitiva en A. Por hipótesis del teorema, tenemos:

$$x\, R_1\, x \,\&\, x\, R_2\, x,$$

de donde

$$x\, R_1 \cap R_2\, x,$$

lo que prueba que $R_1 \cap R_2$ es reflexiva en A.

Supongamos ahora que tenemos ahora para cualesquiera elementos x, y y z de A:

(1) $$x\, R_1 \cap R_2\, y \,\&\, y\, R_1 \cap R_2\, z,$$

lo que es equivalente a

(2) $$xR_1y \,\&\, yR_1z \,\&\, xR_2y \,\&\, yR_2z.$$

De (2) y del hecho de que, por hipótesis, R_1 y R_2 son transitivos en A, inferimos:

(3) $$xR_1z \,\&\, xR_2z,$$

lo que es equivalente a

$$x\, R_1 \cap R_2\, z,$$

y así, $R_1 \cap R_2$ es transitiva en A. Q.E.D.

Si la teoría de las cuasi-ordenaciones se formalizara en lógica de primer orden, necesitaríamos entonces formular el Teorema 1 en forma metamatemática.

TEOREMA 1′. *Si $\langle A, R_1\rangle$ y $\langle A, R_2\rangle$ son modelos que satisfacen la teoría de las cuasi-ordenaciones, entonces $\langle A, R_1 \cap R_2\rangle$ es también uno de esos modelos.*

Esencialmente, la prueba del Teorema 1′ es justamente como la del Teorema 1 y requeriría la misma estructura sinforemática. La ventaja del Teorema 1 sobre el Teorema 1′ es que la prueba del Teorema 1 no requiere ir y venir de una teoría formalizada en lógica de predicados de primer orden a una estructura metamatemática que incluye todo el aparato de la teoría de los conjuntos.

En la demostración informal del Teorema 1 se ha usado en forma intuitiva y un tanto circunstancial propiedades de conjuntos ya conocidas por los Caps. 9 a 11. Es importante repetir la observación hecha en la primera sección de este capítulo; al axiomatizar teorías

particulares *dentro* de la teoría de los conjuntos es posible para casi todos los fines matemáticos, proceder sin tener explícitamente disponible una axiomatización de la propia teoría de los conjuntos. Se seguirá esta práctica en aparatos subsecuentes de este capítulo y es la que debe adoptarse al efectuar varios de los ejercicios. Cuando se axiomatiza una teoría definiendo un predicado sinforemático, se establece la independencia de los axiomas relativos a las nociones primitivas, por la clase de métodos anteriormente descritos y aplicados. Los axiomas se listan en el definiente de la definición del predicado dado y se demuestra que es independiente un axioma dado exhibiendo una entidad sinforemática que satisface el predicado definido por el definiente original, menos el axioma dado; pero que no satisface la totalidad del definiente. Intuitivamente, esto equivale a encontrar un modelo que satisfaga todos, excepto el axioma dado. Es necesario advertir que la estructura puramente sinforemática de las entidades que satisfacen un predicado no está caracterizada en los axiomas propiamente dichos, sino que se enuncia en uno de los tres lugares siguientes:

(i) El texto del definiente que precede inmediatamente a los axiomas (véase la Definición A).
(ii) La hipótesis de la definición, si ésta es de forma condicional (véase la Definición C).
(iii) Informalmente, en la discusión que precede a la definición (véanse las Definiciones D y E).

Desde el punto de vista de la teoría sistemática de la definición (i) es superior, pero desde el punto de vista de la brevedad y la elegancia, es preferible el (iii). Cualquiera alternativa que se adopte, la estructura sinforemática de los modelos de una teoría no se enuncia en las oraciones llamadas *axiomas*.[7] Esto tiene la consecuencia deseable de que las demostraciones de independiencia de los axiomas de una teoría dada hagan todas uso de la misma clase general de entidades sinforemáticas. Por ejemplo, cada uno de los modelos que demuestran la independencia de uno de los axiomas para grupos de la Definición E debe ser un álgebra, esto es, un par ordenado cuyo primer miembro sea un conjunto A y cuyo segundo miembro sea una operación binaria del producto cartesiano $A \times A$ a A.

Puesto que algunas de las definiciones favorecidas en la anterior discusión de alternativas definen predicados de una posición, se necesita alguna aclaración sobre el estado de las nociones primiti-

[7] Cada axioma debería decir algo intuitivamente significativo sobre la teoría. La aserción, por ejemplo, de que una noción primitiva es una relación binaria, difícilmente satisface esta exigencia y, por lo mismo, no debería ser considerado como axioma.

vas. Intuitivamente, las nociones primitivas de una teoría son justamente los conjuntos, relaciones y operaciones que son miembros de un n-tuple ordenado que satisface el predicado dado. Así, la teoría de las cuasi-ordenaciones y la teoría de los grupos, tal como se formulan en este apartado, se basan en dos nociones primitivas.[8] Infortunadamente,cierta ambigüedad rodea el uso del término 'noción primitiva'. Por ejemplo, considérese el predicado 'es un grupo'. Una infinidad de pares ordenados $\langle A, \circ \rangle$ satisfacen este predicado, pero no deseamos inferir del mismo que haya una infinidad de conjuntos y operaciones binarias que son nociones primitivas de la teoría de los grupos. Más bien deseamos decir que (con frecuencia a la definición E) hay exactamente dos nociones primitivas. Podemos convertir en objetos matemáticos definidos las nociones primitivas de una teoría, haciéndolos relativos a un modelo de la teoría. Así, decimos que cualquier modelo de la teoría de los grupos tiene exactamente dos nociones primitivas, a saber, un conjunto y una operación binaria. La prueba de independencia de las nociones primitivas, se limita entonces a demostrar que las nociones primitivas de un modelo arbitrario de la teoría no pueden ser definidas una en función de la otra. Cuando no es posible ninguna confusión nos referiremos simplemente a las nociones primitivas de una *teoría,* pero en todos los casos será perfectamente obvio cómo puede reemplazarse ese lenguaje por la expresión más correcta 'nociones primitivas de un modelo de una teoría'.

Si la formalización de una teoría en lógica de predicados de primer orden se compara con su axiomatización mediante definición de un predicado sinforemático, se notará que casi siempre este último requiere una noción primitiva más, que símbolos primitivos la primera. A saber, allí se encuentra agregada la noción primitiva de un conjunto correspondiente al dominio de un modelo de la teoría formalizada. Por ejemplo, el conjunto A en un grupo $\langle A, \circ \rangle$ no corresponde a ningún símbolo primitivo en la axiomatización correspondiente dada en el §5.2. Las razones para la noción primitiva adicional resultan claras si reflexionamos que cuando los axiomas para grupos se formalizan en lógica de predicados de primer orden, se entiende que el alcance de las variables se extiende a todos los elementos del grupo; pero en la estructura sinforemática de este apartado, el alcance de las variables se extiende a todos los individuos y conjuntos y se necesita la nueva noción primitiva para "relativizar" las variables, de modo que su alcance se extienda a los

[8] Es costumbre referirse a las *nociones primitivas* y no a los símbolos primitivos de una teoría axiomatizada por definición de un predicado sinforemático, pues (hablando literalmente) en la definición del predicado adecuado no se agregan nuevos símbolos primitivos al lenguaje de la teoría de los conjuntos.

elementos del grupo. Así, en la teoría formalizada, tenemos simplemente:

(1) $$x \circ (y \circ z) = (x \circ y) \circ z,$$

pero aquí (1) es reemplazado por:

$$x \in A \,\&\, y \in A \,\&\, z \in A \rightarrow x \circ (y \circ z) = (x \circ y) \circ z,$$

donde $\langle A, \circ \rangle$ es un grupo.

La forma exacta en que se requiere que el conjunto básico de elementos se relacione con las otras nociones primitivas, afecta su independencia como noción primitiva y la verdad o falsedad de ciertos enunciados simples. Por ejemplo, sobre la base de la Definición E, para cualquier grupo $\langle A, \circ \rangle$ el conjunto A puede definirse como el conjunto cuyo producto cartesiano es el dominio de definición de la operación binaria $\circ$. Mediante una ligera reformulación de los requerimientos sinforemáticos generales, puede hacerse el conjunto básico A independiente de la otra noción primitva. Esto puede ilustrarse reformulando la Definición E de grupos. Los dos cambios esenciales consisten en abandonar la exigencia de que $\circ$ sea una operación binaria de $A \times A$ a A y añadir el *axioma de oclusión* de que si x, $y \in A$, entonces $x \circ y \in A$. Una *álgebra* $\langle A, \circ \rangle$ se define ahora simplemente como un conjunto no vacío A y una operación binaria $\circ$, sin imponer ninguna restricción a la relación entre A y $\circ$, y definimos los grupos mediante la

DEFINICIÓN J. *Una álgebra* $\mathfrak{A} = \langle A, \circ \rangle$ *es un grupo si y sólo si para toda x, y y z, de A* .

A1. $x \circ y \in A$.
A2. $x \circ (y \circ z) = (x \circ y) \circ z$.
A3. *Hay una w en A tal que*

$$x = y \circ w.$$

A4. *Hay una w en A tal que*

$$x = w \circ y.$$

Los siguientes dos grupos que satisfacen la Definición J demuestran que A es independiente. Sea $+$ el conjunto de triples ordenados de enteros $\langle x, y, z \rangle$, tales que $x + y = z$. Sea A_1 el conjunto de enteros y sea A_2 el conjunto de enteros pares, incluido el cero. Entonces $\langle A_1 +\rangle$ y $\langle A_2 +\rangle$ son ambos grupos en el sentido de la Definición J

y por aplicación del principio de Padoa (véase el Cap. 8) está claro que estos dos grupos establecen la independencia de *A* como una noción primitiva. No obstante, la obtención de independencia de las nociones primitivas no es suficiente para justificar el uso de la Definición J en vez de la Definición E. De hecho, muchos algebristas prefieren E. En el apartado siguiente se presentarán muchas razones. Por otra parte, en los ejercicios se menciona una ventaja de J. Nótese que hemos definido las cuasi-ordenaciones de una manera análoga a J, más bien que a E.

En muchos contextos (se da un ejemplo de uno de ellos en el apartado siguiente) es conveniente hacer que un grupo sea un cuádruple ordenado, en vez de un par ordenado. E introducimos álgebras que son cuádruples ordenados $\langle A, \circ, ^{-1}, e\rangle$, donde $\circ$ es una operación binaria de $A \times A$ a A, $^{-1}$ es una operación unaria de A a A, y e es un elemento de A. Correspondiendo a los axiomas para grupos dados en el texto de §5.1, tenemos entonces:

Definición K: *Una álgebra* $\mathfrak{A} = \langle A, \circ, ^{-1}, e\rangle$ *es un grupo si y sólo si, para toda x, y y z de A*

A1. $x \circ (y \circ z) = (x \circ y) \circ z$.
A2. $x \circ e = x$.
A3. $x \circ x^{-1} = e$.

A primera vista podría parecer que la adición de la Definición K introduce una ambigüedad en nuestro uso del predicado 'es un grupo'. Sin embargo, nótese que tanto la Definición E como la Definición K son definiciones condicionales disfrazadas, particular hecho explícito por la Definición F. Las hipótesis apropiadas de las dos deficiones son mutuamente exclusivas, de manera que cualquiera entidad sinforemática que satisface una, no puede satisfacer otra. En particular, la hipótesis de la Definición E es que $\mathfrak{A}$ es un par ordenado $\langle A, \circ\rangle$, y la hipótesis de la Definición K, que $\mathfrak{A}$ es un cuádruple ordenado $\langle A, \circ, ^{-1}, e\rangle$. El siguiente teorema correlaciona las dos definiciones en forma exacta, del cual hacemos uso en el apartado siguiente. La prueba se deja como ejercicio. (La palabra álgebra en la hipótesis del teorema se usa en el sentido definido justamente antes de la Definición E.)

Teorema 2. *Si* $\langle A, \circ\rangle$ *es una álgebra,* $\langle A, \circ\rangle$ *es un grupo (en el sentido de la Definición E) si y sólo si hay una operación unaria* $^{-1}$ *de A a A y de un elemento e de A tal que* $\langle A, \circ, ^{-1}, e\rangle$ *es un grupo (en el sentido de la Definición K).*

En los §12.4 y §12.5 pasamos a ver dos ejemplos de axiomatización que siguen las ideas expuestas en el presente apartado. El

primer ejemplo se refiere a probabilidad y el segundo a mecánica. Ninguno de los dos tiene una formalización natural y sencilla en la lógica de predicados de primer orden. Consecuentemente, se pretende que cada uno ejemplifique la potencia y flexibilidad relativas del procedimiento axiomático, que consiste en definir un predicado sinforemático adecuado. Antes de pasar a estos dos ejemplos, se introduce en el apartado siguiente la importante noción de isomorfismo de modelos para una teoría.

EJERCICIOS

1. Axiomatícese la teoría de las álgebras booleanas discutidas en el §9.9 (véase particularmente el Ejercicio 9), definiendo los predicados sinforemáticos adecuados,

   (a) en el estilo de la Definición A;
   (b) en el estilo de la Definición E;
   (c) en el estilo de la Definición J (hay que cerciorarse de que no se incluyen axiomas de oclusión en el caso).

2. Axiomatícese la teoría de la medición de masas discutida en el Ejercicio 3 del § 4.5, definiendo el predicado sinforemático apropiado.

3. Usando la Definición A, decídase qué es verdadero de lo que sigue. Si es falso, dese un contraejemplo.

   (a) Si $\langle A, R_1\rangle$ y $\langle A, R_2\rangle$ son cuasi-ordenaciones, entonces $\langle A, R_1 \cup R_2\rangle$ es una cuasi-ordenación.
   (b) Si $\langle A, R_1\rangle$ y $\langle A, R_2\rangle$ son cuasi-ordenaciones, entonces $\langle A, R_1 \sim R_2\rangle$ es una cuasi-ordenación.
   (c) Si $\langle A, R_1\rangle$ y $\langle A, R_2\rangle$ son cuasi-ordenaciones, entonces $\langle A, R_1/R_2\rangle$ es una cuasi-ordenación.
   (d) Si $\langle A_1, R\rangle$ y $\langle A_2, R\rangle$ son cuasi-ordenaciones, entonces $\langle A_1 \cap A_2, R\rangle$ es una cuasi-ordenación.
   (e) Si $\langle A_1, R\rangle$ y $\langle A_2, R\rangle$ son cuasi-ordenaciones, entonces $\langle A_1 \cup A_2, R\rangle$ es una cuasi-ordenación.
   (f) Si $\langle A_1, R_1\rangle$ y $\langle A_2, R_2\rangle$ son cuasi-ordenaciones, entonces $\langle A_1 \cap A_2, R_1 \cap R_2\rangle$ es una cuasi-ordenación.
   (g) Si $\langle A, R_1\rangle$ y $\langle A, R_2\rangle$ son cuasi-ordenaciones, entonces $\langle A \times A, R_1 \times R_2\rangle$ es una cuasi-ordenación.

4. Uno de los problemas fundamentales de la economía del bienestar y de la teoría política es el del modo óptimo de agregación de preferencias individuales para determinar preferencias de la unidad social. Para ambas unidades, la individual y la social, definimos:[9] *$\mathfrak{A}$ es un patrón de preferencia débil si y sólo si hay un conjunto A y relaciones binarias P e I, tales que* $\mathfrak{A} = \langle A, P, I\rangle$ *y*

[9] Usamos el adjetivo 'débil' porque los axiomas no dicen mucho sobre la teoría de la preferencia. Para una extensa discusión de este problema en el estilo de este ejercicio, véase K. J. Arrow, *Social Choice and Individual Values*, Nueva York, 1951.

AXIOMA 1. *P es transitivo en A.*

AXIOMA 2. *I es transitivo en A.*

AXIOMA 3. *Para cualquiera x y y en A, exactamente uno de los siguientes: xPy, yPx, xIy.*

(Para una interpretación intuitiva véase el Ejercicio 6 del § 10.5.)

Consideremos ahora una unidad social que no consiste solamente en dos individuos con patrones de preferencia débil $\langle A, P_1, I_1\rangle$ y $\langle A_2, P_2, I_2\rangle$ (para algún conjunto fijo de alternativas confrontadas con la unidad).

(a) ¿Es $\langle A, P_1 \cap P_2, I_1 \cap I_2\rangle$ un patrón de preferencia débil (para la unidad)? Esto es, ¿se satisfacen los tres axiomas?

(b) Defínase para $x, y \in A$:

$$x\,P\,y \leftrightarrow [x\,P_1\,y\;\&\;(x\,P_2\,y \vee x\,I_2\,y)] \vee [x\,P_2\,y\;\&\;(x\,P_1\,y \vee x\,I_1\,y)],$$

$$x\,I\,y \leftrightarrow -(x\,P\,y)\;\&\;-(y\,P\,x).$$

¿Es $\langle A, P, I\rangle$ un patrón de preferencia débil para la unidad?

(c) Describa formalmente, en la manera de (b), el método de agregación que piense que más estrechamente corresponda al método de voto de mayoría para una unidad social de tres individuos.

5. Siguiendo los desarrollos del Ejercicio 5 del § 5.2, no es difícil justificar la adición de las dos equivalencias definicionales siguientes a los axiomas para grupos de la Definición J:

$$e = y \leftrightarrow (x)(x \in A \rightarrow x \circ y = x \;\&\; y \circ x = x),$$

$$x^{-1} = y \leftrightarrow x \circ y = e.$$

(En otras palabras, introducimos por medio de estas definiciones el elemento de identidad y la operación inversa de un grupo.) Definimos ahora: *si* $\mathfrak{A} = \langle A, \circ\rangle$ *es un grupo y* $B \subseteq A$, *entonces* $\langle B, \circ\rangle$ *es un subgrupo de* $\mathfrak{A}$ *si y sólo si para cada x y y en B*

(i) $x \circ y \in B$

(ii) $e \in B$

(iii) $x^{-1} \in B$.

(Nótese que la definición de subgrupo debe complicarse si se usa la Definición E para grupos, pues la operación $\circ$ no es una operación de $B \times B$ a B.)

(a) Pruébese que si $B \neq \Lambda$ y $B \subseteq A$, entonces $\langle B, \circ\rangle$ es un subgrupo de $\mathfrak{A}$ si y sólo si para cada x y y de B, $x \circ y \in B$ y $x^{-1} \in B$.

(b) Pruébese que si $\langle B_1, \circ\rangle$ y $\langle B_2, \circ\rangle$ son subgrupos de $\mathfrak{A}$, entonces $\langle B_1 \cap B_2, \circ\rangle$ es un subgrupo de $\mathfrak{A}$.

(c) Si $\langle B_1, \circ\rangle$ y $\langle B_2, \circ\rangle$ son subgrupos de $\mathfrak{A}$, ¿es $\langle B_1 \cup B_2, \circ\rangle$ un subgrupo de $\mathfrak{A}$?

(d) Para $B_1 \subseteq A$ y $B_2 \subseteq A$, defínase

$$B_1 \circ B_2 = C \leftrightarrow (z)(z \in C \leftrightarrow (\exists x)(\exists y)(x \in B_1 \;\&\; y \in B_2 \;\&\; x \circ y = z).$$

Pruébese que si $\langle B_1, \circ\rangle$ y $\langle B_2, \circ\rangle$ son subgrupos de $\mathfrak{A}$, entonces $\langle B_1 \circ B_2, \circ\rangle$ es un subgrupo de $\mathfrak{A}$, si y sólo si $B_1 \circ B_2 = B_2 \circ B_1$.

6. Haciendo referencia al ejercicio precedente, defínase la noción de subgrupo cuando se usa la Definición E para grupos (y así la noción de una álgebra corresponde a la definida inmediatamente antes de la Definición E).

Defina también la noción de subgrupo, para grupos $\langle A, \circ, ^{-1}, e\rangle$ que satisfacen la Definición K.

7. Demuéstrese el Teorema 2.

§12.3. Isomorfismo de modelos para una teoría. La separación de la caracterización puramente sinforemática de la estructura de los modelos de una teoría, de los axiomas propiamente dichos, es significativa para definir ciertas importantes nociones relativas a los modelos de una teoría. Por ejemplo, se dice a menudo que la noción de que dos grupos son isomórficos está libre de axiomas, puesto que la definición de isomorfismo de grupos no depende de ninguno de los axiomas. De hecho, la definición es realmente para álgebras isomórficas y se aplica a álgebras que no son grupos. Un rápido examen de la definición, que pasamos a enunciar, comprueba estas observaciones. (En la definición, se entiende que una álgebra es un par ordenado $\langle A, \circ\rangle$, tal que $\circ$ es una operación binaria de $A \times A$ a A, que es el primero de los sentidos de álgebra introducido en la sección anterior.)

DEFINICIÓN 1. *Una álgebra* $\mathfrak{A} = \langle A, \circ\rangle$ *es isomórfica con respecto de una álgebra* $\mathfrak{A}' = \langle A', \circ'\rangle$ *si y sólo si hay una función f tal que*

(i) $D(f) = A \ \& \ \mathscr{R}(f) = A'$,
(ii) f es una función uno-uno,
(iii) si $x, y \in A$, entonces

$$f(x \circ y) = f(x) \circ' f(y).$$

La idea intuitiva es, por supuesto, que dos álgebras son isomórficas sólo cuando tienen la misma estructura. Las propiedades requeridas de la función f hacen precisa la idea de *misma estructura.* La relación de isomorfismo entre álgebras es una relación de equivalencia, esto es, reflexiva, simétrica y transitiva. (Dejamos la demostración como ejercicio.) Un ejemplo de dos grupos distintos que son isomórficos es suministrado por lo que sigue. Sea A el conjunto de enteros, sea A' el conjunto de enteros pares y sea $+'$ la adición de enteros pares (donde se cuenta cero como entero par). Entonces $\langle A, +\rangle$ es isomórfico con respecto de $\langle A', +'\rangle$, pues es fácil encontrar una función f con propiedades de (i) a (iii). Sea f una función tal que, para cualquier entero n

$$f(n) = 2n.$$

Verificamos entonces en forma inmediata que $D(f) = A$, $\mathscr{R}(f) = A'$ y f es una función uno-uno. Además, para dos enteros cualesquiera m y n

$$f(m+n) = 2(m+n) = 2m +' 2n = f(m) +' f(n),$$

que verifica (iii).

Si una álgebra se define como un par ordenado $\langle A, \circ \rangle$ donde la operación $\circ$ necesita no tener conexión con el conjunto A, la Definición 1 debe ser modificada, pues $x \circ y$ puede no estar en el dominio de f. Haciendo hincapié en que las operaciones $\circ$ y $\circ'$ son conjuntos de triples ordenados, reemplazamos (iii) de la Definición 1 por

(iii') Si $x, y, z \in A$, entonces

$$\langle x, y, z \rangle \in \circ \leftrightarrow \langle f(x), f(y), f(z) \rangle \in \circ'.$$

Podemos dar a (iii') la forma

(iii'') Si $x, y, z \in A$, entonces

$$x \circ y = z \leftrightarrow f(x) \circ' f(y) = f(z).$$

Cuando se trata de álgebras en este segundo sentido debe tenerse cuidado de determinar cuándo la ejecución de la operación $\circ$ en dos elementos del conjunto A da como resultado un elemento de A. Además, sólo en las circunstancias más triviales puede determinarse si dos álgebras son idénticas, puesto que las operaciones binarias de dos álgebras cualesquiera pueden tener cualquier clase de estructura fuera del conjunto $A \times A \times A$. En los ejercicios del final de este apartado se consideran algunas álgebras $\langle A, \circ \rangle$ en el primer sentido, esto es, donde la operación $\circ$ es de $A \times A$ a A.

Una definición general satisfactoria de isomorfismo para dos entidades sinforemáticas de cualquier clase es difícil, si no es que imposible de formular. La práctica matemática típica es la de formular una definición separada para cada clase general de n-tuple ordenado. Cuando los n-tuples son complicados, como en el caso de modelos para la teoría de la mecánica de puntos materiales, es a veces difícil decidir exactamente qué se da a entender con dos modelos isomórficos; pero para álgebras o simples estructuras de relación es más clara la selección de las condiciones apropiadas. Para ilustrar la condición impuesta a las relaciones binarias, enunciamos la definición de isomorfismo para estructuras simples de relación.

DEFINICIÓN 2. *Una estructura simple de relación* $\langle A, R \rangle$ *es isomórfica con respecto de una estructura simple de relación* $\langle B, S \rangle$, *si y sólo si hay una función* f *tal que*

(i) $D(f) = A \;\&\; \mathscr{R}(f) = B$,
(ii) f es una función uno-uno,
(iii) Si $x, y \in A$, entonces $xRy \leftrightarrow f(x)\,S\,f(y)$.

(En lugar de '$\langle B, S\rangle$' podríamos igualmente haber escrito '$\langle A', R'\rangle$' con cambios correspondientes en (i) a (iii).)

Usando esta definición, la estructura de relación

$$\langle\{1, 2\}, \{\langle 1, 2\rangle, \langle 2, 2\rangle\}\rangle$$

es isomórfica con respecto de la estructura de relación

$$\langle\{3, 4\}, \{\langle 4, 3\rangle, \langle 3, 3\rangle\}\rangle,$$

pues podemos tomar como una función apropiada de f:

$$f(1) = 4$$

$$f(2) = 3.$$

La noción de isomorfismo tiene importantes aplicaciones en todos los dominios de las matemáticas modernas. Cuando se da la situación especial de que son isomórficos dos modelos cualesquiera de una teoría, entonces se dice que la teoría es *categórica*.[10] Pueden encontrarse sencillos contraejemplos para demostrar que ni la teoría de las cuasi-ordenaciones ni la teoría de los grupos es categórica. Por ejemplo, en el caso de las cuasi-ordenaciones podemos tomar dos cuasi-ordenaciones $\langle A, R\rangle$ y $\langle B, S\rangle$ tales que el conjunto A tiene dos elementos y el conjunto B tres elementos. No hay, pues, una función uno-uno cuyo dominio sea A y su alcance B. Esto es trivial. Es un poco más interesante exhibir dos cuasi-ordenaciones $\langle A, R\rangle$ y $\langle B, S\rangle$, tales que A y B tengan el mismo número de elementos y, no obstante, $\langle A, R\rangle$ no sea isomórfico con respecto de $\langle B, S\rangle$. Sea

$$A = \{1, 2\}$$

$$R = \{\langle 1, 1\rangle, \langle 2, 2\rangle, \langle 1, 2\rangle\}$$

$$B = \{3, 4\}$$

$$S = \{\langle 3, 3\rangle, \langle 4, 4\rangle\}.$$

Entonces, si existiera la f apropiada para el isomorfismo de $\langle A, R\rangle$ y $\langle B, S\rangle$, de

$$1\,R\,2$$

[10] Esta noción surge (1904) con el matemático americano Oswal Veblen. Puede ser de cierto interés para los filósofos saber que la palabra 'categórico' fue sugerida a Veblen por John Dewey (véase *Actas de la Sociedad Matemática Americana*, Vol. 5 (1904). Pág. 346).

inferiríamos

$$f(1)\, S\, f(2).$$

Puesto que f es una función uno-uno, $f(1) \neq f(2)$, pero la relación S no es válida entre dos elementos distintos. Consecuentemente, no puede existir esa función f.

En los ejercicios se dan algunos ejemplos sencillos de teorías categóricas.[11] Cuando una teoría no es categórica, un problema importante es el de descubrir si puede encontrarse un subconjunto interesante de modelos de la teoría, tal que cualquier modelo de la teoría sea isomórfico con respecto de algún miembro de este subconjunto. Encontrar ese destacado subconjunto de modelos de una teoría y demostrar que tiene la propiedad indicada es demostrar un *teorema de representación* de la teoría. Ese teorema puede demostrarse por grupos; a saber, cada grupo es isomórfico con respecto de un *grupo de transformaciones*. Hablando en términos generales, un grupo de transformaciones puede ser descrito como sigue. Sea M un conjunto no vacío y B un conjunto de funciones uno-uno cuyos dominios y alcances son M (esa función se llama una *transformación* sobre M). Entonces, si B es adecuadamente escogido $\langle B, \circ \rangle$ será un *grupo* de transformaciones, donde $\circ$ es composición de funciones. Así, si $M = \{1, 2\}$, y B = el conjunto de todas las transformaciones sobre M, entonces hay dos funciones f_1 y f_2 en B:

$$f_1(1) = 1 \;\&\; f_1(2) = 2,$$
$$f_2(1) = 2 \;\&\; f_2(2) = 1.$$

La función f_1 es la transformación de *identidad*, puesto que traza el mapa de cada elemento M dentro de sí misma. Notamos también que tanto f_1 como f_2 son idénticas a sus inversas, esto es,

$$f_1^{-1} = f_1$$
$$f_2^{-1} = f_2.$$

Podemos ahora comprobar si $\langle B, \circ \rangle$ es un grupo de transformaciones (donde $\circ$ es composición de funciones). Al hacer esa comprobación es conveniente usar el Teorema 2 del No. 12.2; esto es, demostramos que el álgebra $\mathfrak{B} = \langle B, \circ \rangle$ es un grupo, demostrando que hay una operación inversa $^{-1}$ y un elemento de identidad e, tal que $\langle B, \circ, ^{-1}, e \rangle$ es un grupo. Vemos inmediatamente que el elemento de identidad de $\mathfrak{B}$ es la identidad de transformación f_1, puesto que

[11] La teoría de los números reales y la teoría de los enteros positivos son probablemente los dos ejemplos más importantes de teorías categóricas.

$$f_1 \circ f_2 = f_2 \circ f_1 = f_2.$$

Y el elemento inverso de grupo de una función en B es justamente la función inversa, de donde, inmediatamente

$$f_1 \circ f_1^{-1} = f_1^{-1} \circ f_1 = f_1$$
$$f_2 \circ f_2^{-1} = f_2^{-1} \circ f_2 = f_1.$$

A continuación probamos el teorema de representación mencionado.[12]

TEOREMA 2. *Todo grupo es isomórfico con respecto de un grupo de transformaciones.*

PRUEBA. Sea $\mathfrak{A} = \langle A, \circ \rangle$ un grupo arbitrario. Para cada x de A definimos una función f_x que traza el mapa de A en A como sigue: para cada y de A

$$(1) \qquad f_x(y) = x \circ y.$$

De la ley de cancelación a mano izquierda para grupos se sigue (§ 5.2, Ejercicio 4, Teorema 5), que f_x es una función uno-uno. Porque supóngase que no lo fuera. Entonces habría elementos y y y' en A, tales que $y \neq y'$ y

$$f_x(y) = f_x(y');$$

pero entonces por (1)

$$x \circ y = x \circ y',$$

y por la ley de cancelación

$$y = y',$$

lo cual es absurdo.

Del Axioma A2 de la Definición E se sigue que para cualquiera z de A hay una y en A tal que

$$f_x(y) = z.$$

Así, el alcance de f_x es A, de donde f_x es una transformación.

Demostremos ahora que el conjunto de todas esas transformaciones f_x forma un grupo con respecto de la operación de composición

[12] Los lectores interesados en un ulterior desarrollo de la teoría de grupos y tópicos correlativos del álgebra moderna encontrarán útil la obra de G. Birkhoff y S. MacLane, *A Survey of Modern Algebra*, edición revisada, Nueva York, 1953.

de funciones. (Para evitar confusiones entre la operación de composición de funciones y la operación de grupo de $\mathfrak{A}$, denotamos composición de funciones por yuxtaposición en esta demostración.) Usamos el Teorema 2 del § 12.2. Sea e el elemento de identidad de $\mathfrak{A}$; entonces f_e es la transformación de identidad sobre A, puesto que

$$f_e(x) = e \circ x = x.$$

De donde,

$$f_x f_e = f_e f_x = f_x.$$

Además, puesto que para cualquier x de A, f_x es 1 a 1, la función inversa f_x^{-1} existe, y tenemos inmediatamente

$$f_x f_x^{-1} = f_x^{-1} f_x = f_e.$$

Finalmente, para completar la prueba de nuestro teorema, necesitamos demostrar que el grupo de transformaciones sobre A que hemos definido es isomórfico con respecto de $\mathfrak{A}$. Por esto, necesitamos primero observar que para cualquier x y y de A

$$f_x f_y = f_{x \circ y},$$

puesto que para cualquier z de A

(2) $$f_x f_y(z) = f_x(y \circ z) = x \circ (y \circ z) = (x \circ y) \circ z = f_{x \circ y}(z).$$

Ahora para la función que demuestra el isomorfismo de los dos grupos, usamos, como sería de esperarse, una función φ, tal que para cualquier x de A

$$\varphi(x) = f_x.$$

Obviamente, el dominio y el alcance de φ son lo que deberían ser. Para ver que φ es una función uno-uno, notamos que si $x \neq y$, entonces

$$f_x(e) \neq f_y(e),$$

pues si $f_x = fy$ cuando $x \neq y$, entonces

$$x = x \circ e = f_x(e) = f_y(e) = y \circ e = y.$$

Para establecer (iii) de la Definición 1, usamos (2), a fin de inferir que

$$\varphi(x \circ y) = f_{x \circ y} = f_x f_y = \varphi(x)\varphi(y). \qquad \text{Q.E.D.}$$

La noción de un teorema de representación y, en consecuencia, la noción de isomorfismo, tiene importantes aplicaciones en la filosofía de la ciencia. La meta primaria de la teoría de la medición, por ejemplo, es demostrar en forma precisa cómo pasar de observaciones cualitativas ('esta varilla es más larga que ésta', 'el platillo izquierdo de la balanza está más alto que el derecho') a las aserciones cuantitativas que se necesitan en las ciencias empíricas, ('la longitud de esta varilla es de 7.2 cm'; 'la masa de esta muestra química es de 5.4 g'). En otras palabras, la teoría de las mediciones debería suministrar un análisis exacto de cómo podemos inferir aserciones cuantitativas, de observaciones fundamentalmente cualitativas. Ese análisis lo suministra la axiomatización de álgebras apropiadas de operaciones y relaciones experimentalmente realizables. El Ejercicio 9 del § 4.5 constituyó un ejemplo parcial de medición de masas. Dada una teoría axiomatizada de medición de alguna cantidad empírica, la tarea matemática consiste en demostrar un teorema de representación para modelos de la teoría, que establece, hablando en términos generales, que cualquier modelo es isomórfico con respecto de algún modelo numérico de la teoría. La existencia de este isomorfismo justifica la aplicación de números a las cosas. No podemos literalmente tomar un número y "aplicarlo" a un objeto físico, digamos. Lo que podemos hacer es demostrar que la estructura de un conjunto de fenómenos, en ciertas operaciones y relaciones empíricas, es *igual que* la estructura de algún conjunto de números, en ciertas operaciones y relaciones aritméticas. La definición de isomorfismo en el contexto dado hace adquirir precisión a la idea intuitiva de *igual estructura*, como ya se ha hecho notar. La gran importancia de encontrar ese isomorfismo de estructuras estriba en que podemos entonces usar todo nuestro conocimiento de los métodos de cálculo, como se aplican a la estructura aritmética, para inferir hechos sobre la estructura empírica isomórfica.

Desgraciadamente, la mayor parte de las pruebas de los teoremas de representación en la teoría de la medición son demasiado largas para incluirse en el texto o en los ejercicios. Para ilustrar los métodos, consideraremos un conjunto muy sencillo de axiomas para medir diferencias percibidas o sentidas en varias clases de fenómenos, como diferencias de intensidad o de tono, de un conjunto de sonidos, diferencia en la intensidad de un dolor o de un conjunto

de estímulos, diferencia en la brillantez visual de un conjunto de estímulos de color y diferencias de valor (o utilidad) de un conjunto de bienes económicos. Los axiomas están basados en tres nociones primitivas: el conjunto A de objetos o estímulos, la relación binaria P que representa la ordenación de los estímulos y la relación cuaternaria E que representa igualdad de diferencia de pares de estímulos.

Supóngase, por ejemplo, que deseamos desarrollar una escala numérica para el tono. Entonces A sería cierto conjunto de sonidos. Para dos sonidos x y y de A, xPy si y sólo si se juzga que el tono de x es inferior al de y. Para cuatro sonidos x, y, u y v en A, $x, yEu,$ v si y sólo si la diferencia (algebraica) de tono entre x y y se estima igual que la existente entre u y v. Decimos diferencia *algebraica*, puesto que E toma en cuenta la ordenación; esto es, si x, y, u y v fueran números, entonces tendríamos:

$$x, y\, E\, u, v \leftrightarrow x - y = u - v.$$

Hablando en general, nuestro objetivo es el de enunciar axiomas sobre A, P y E, tales que podamos demostrar que existe una función φ con valores reales, definida sobre A, tal que

$$xPy \leftrightarrow \varphi(x) < \varphi(y),$$

$$x, y\, E\, u, v \leftrightarrow \varphi(x) - \varphi(y) = \varphi(u) - \varphi(v).$$

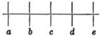
Procedemos ahora a sistematizar nuestras ideas. Teniendo a la vista aplicaciones en psicología, llamamos a cualquier triple ordenado $\langle A, P, E\rangle$, donde A es un conjunto, P una relación binaria y E una relación cuaternaria, una *estructura de diferencias*. Para poner de relieve que la definición apropiada de isomorfismo es libre de axiomas, podemos enunciarla antes de considerar los axiomas de nuestra teoría de medición.

DEFINICIÓN 3. *Una estructura de diferencia* $\langle A, P, E\rangle$ *es isomórfica con respecto de una estructura de diferencia* $\langle A' P', E'\rangle$ *si y sólo si hay una función f tal que*

(i) $D(f) = A$ & $\mathscr{R}(f) = A'$,
(ii) f es una función uno-uno,
(iii) Si $x, y \in A$, entonces $xPy \leftrightarrow f(x)P'f(y)$,
(iv) Si $x, y, u, v \in A$, entonces $x, y\,E\,u, v \leftrightarrow f(x), f(y)E'f(u), f(v)$.

Esta definición es simplemente una extensión en forma natural de la Definición 2. Pasamos ahora a los axiomas que debe satisfacer

una estructura de diferencias, a fin de ser una estructura de diferencias iguales. La idea intuitiva de la axiomatización es requerir que los estímulos se ordenen conforme al tono, intensidad de brillantez visual, o alguna otra característica y a continuación exigir que dos estímulos adyacentes cualesquiera de la ordenación tengan la misma diferencia de intensidad que cualesquiera otros dos de esos estímulos adyacentes. Además, nos restringimos a conjuntos *finitos* de estímulos. Un conjunto de cinco estímulos tendría que ser arreglado como



para ser una estructura de diferencias iguales. Obviamente, podemos asignar el número 1 a *a*, 2 a *b*, etc. Nuestros axiomas enuncian suficientes condiciones para garantizar que es siempre posible esa asignación.

Antes de enunciar los axiomas, es necesario introducir una función definida que sea de uso cómodo tanto en los axiomas como en los teoremas. Además, la introducción de esta noción definida puede servir para ilustrar algunos puntos pertinentes relativos a la forma adecuada de definición de las nociones definidas.

DEFINICIÓN 4A. *Si $x, y \in A$, entonces*

$$xJy \leftrightarrow [xPy \,\&\, (z)(z \in A \,\&\, xPz \rightarrow y = z \vee yPz)].$$

La interpretación intuitiva de la relación binaria *J es* que *xJy* si y sólo si *y* es el *sucesor único inmediato* de *x* con respecto de la relación P. Obviamente la Definición 4A no es una definición adecuada en la teoría general de los conjuntos, pues el definiente contiene las variables libres 'A' y 'P', que no aparecen en el definiendo. La definición sería apropiada solamente si 'A' y 'P' fueran constantes y no variables. Una definición adecuada es:

DEFINICIÓN 4B. *Si ⟨A, P, E⟩ es una estructura de diferencia y si $x, y \in A$, entonces*

$$xJ(A, P)y \leftrightarrow [xPy \,\&\, (z)(z \in A \,\&\, xPz \rightarrow y = z \vee yPz)].$$

También, puesto que la variable 'E' no ocurre en el consecuente de 4B, podemos escribir:

DEFINICIÓN 4C. *Si hay una E tal que ⟨A, P, E⟩ es una estructura y si $x, y \in A$, entonces*

$$xJ(A, P)y \leftrightarrow [xPy \,\&\, (z)(z \in A \,\&\, xPz \rightarrow y = z \vee yPz)].$$

Pero puesto que para cualquier conjunto A y relación binaria P hay una E tal que $\langle A, E, P\rangle$ es una estructura de diferencia, podemos también simplificar la definición y hacer referencia solamente a estructuras simples de relación $\langle A, P\rangle$.

DEFINICIÓN 4D. *Si $\langle A, P\rangle$ es una estructura simple de relación y si $x, y \in A$, entonces*

$$xJ(A, P)y \leftrightarrow [xPy \,\&\, (z)(z \in A \,\&\, xPz \rightarrow y = z \vee yPz)].$$

Para 4B, 4C y 4D se satisface la regla de las definiciones condicionales de símbolos de relación y 'J' es un símbolo cuaternario de relación definido de la teoría de los conjuntos. Pero este resultado parece ligeramente artificial. Esperamos que 'J' tenga la misma categoría que 'A' y 'P', que son variables. Por esta razón, la práctica universal es la de usar 4A y explicar sus limitaciones en la forma siguiente. Básicamente la definición de la relación J es la clase de definición que podríamos añadir como axioma definicional a cualquier lista de axiomas que impusiérâmos a las estructuras de diferencia. Si nuestra teoría de la medición se axiomatizara en lógica de primer orden, entonces 'A' y 'P' serían símbolos primitivos y 'J' sería un símbolo de relación binaria apropiadamente introducido por 4A. Esas definiciones (esto es, 4A) se denominan 'elementales' con respecto de la teoría que se considera y se usan constantemente para evitar la molesta notación requerida por las definiciones de la teoría general de los conjuntos, como 4B, 4C y 4D. Cuando se considera alguna teoría fija, se antepone el prefijo 'E' a definiciones como 4A, para indicar su carácter elemental.[13] La misma marca es necesaria en los teoremas elementales, que son teoremas deducibles de los axiomas de la teoría dada y que son válidos para cualquier modelo de la teoría. (Volveremos a este particular.)

Volvemos a nuestra tarea principal, el desarrollo de la teoría de la medición, para estructuras finitas de diferencias iguales. La relación J se enuncia en el axioma final de la definición básica que ahora enunciamos.[14]

[13] Todas estas observaciones se aplican a la definición del elemento de identidad y a la operación inversa para grupos $\langle A, \circ\rangle$. Así, cuando se usa una definición en la teoría general de los conjuntos, se denota el elemento de identidad o por medio de un símbolo de operación binaria, '$e(A, \circ)$': Si una álgebra $\langle A, \circ\rangle$ es un grupo, entonces

$$e(A, \circ) = y \leftrightarrow (x)(x \in A \rightarrow x \circ y = x).$$

[14] Un conjunto relacionado de axiomas se encuentra en la obra de D. Davidson y P. Suppes, "A Finitistic Axiomatization of Subjective Probabilty and Utility", *Econometrica*, Vol. 24 (1956), Págs. 264 a 275. Un conjunto de axiomas para estructuras de diferencias infinitas se da en P. Suppes y M. Winet, "An Axiomatization of Utility Based on the Notion of Utility Differences", *Management Science*, Vol. 1 (1955), Págs. 259 a 270.

DEFINICIÓN 5. *Una estructura de diferencia* $\langle A, P, E\rangle$ *es una estructura finita de diferencias iguales si y sólo si A es un conjunto finito y para cada x, y, z, v y u de A, se satisfacen los siguientes nuevos axiomas:*

AXIOMA D1. *Si xPy, entonces no yPx.*
AXIOMA D2. *Si xPy & yPz, entonces xPz.*
AXIOMA D3. *Si* $x \neq y$, *entonces xPy o yPx.*
AXIOMA D4. *x, x E y, y.*
AXIOMA D5. *Si x, y E u, v entonces u, v E x, y.*
AXIOMA D6. *Si x, y E u, v, entonces x, u E y, v.*
AXIOMA D7. *Si x y E u, v & u, v E z, w, entonces x, E z, w.*
AXIOMA D8. *Si xPy & x, y E u, v, entonces uPv.*
AXIOMA D9. *Si xJy & uJv, entonces x, y E u, v.*

Recordando la interpretación empírica discutida anteriormente, podemos captar la interpretación intuitiva de cada axioma. Por ejemplo, los primeros tres axiomas requieren solamente que P sea una ordenación estricta simple de A. Los cuatro axiomas siguientes establecen solamente propiedades de E. Por ejemplo, el Axioma D4 dice que la diferencia de tono entre cualquier sonido x y el mismo, es igual a la diferencia de tono existente entre cualquier otro sonido y y el mismo. Sobre la base del conjunto total de axiomas podemos probar que esta diferencia es igual a cero, como era de esperarse. La interpretación numérica de D4 es sólo que

$$x - x = y - y,$$

una trivial verdad aritmética para cualesquiera números x y y. Los dos axiomas finales relacionan entre sí a P y E. El Axioma D9 es el que impone el espaciamiento igual de diferencias a los estímulos. Respecto de la aplicación experimental de estos axiomas, debe advertirse que D9 tiene una categoría diferente a la de los axiomas D1 a D8. Podríamos esperar que un sujeto cuidadoso satisficiera aproximadamente de D1 a D8 para cualquier conjunto de estímulos, con respecto de una característica como el tono o la intensidad lumínica. Pero sería verdaderamente sorprendente encontrar que D9 era satisfecho por un conjunto de estímulos arbitrariamente elegido. El conjunto adecuado de estímulos debe ser cuidadosamente escogido por el experimentador, a fin de encontrar un modelo que satisfaga a D9.[15]

Los teoremas elementales que se han necesitado para demostrar nuestro teorema de representación, relativo a estructuras de diferencias finitas iguales, se listan en los ejercicios del extremo de esta sección. De las observaciones anteriores sobre definiciones elementales debe resultar claro cómo se puede convertir cualquier teorema elemental sobre estructuras de diferencias finitas iguales en un teorema de teoría general de los conjuntos. Por ejemplo, el primer teorema es:

TEOREMA E1. *Si* $x, y, z \in A$ *y si* $x, y\ E\ x, z$, *entonces* $y = z$.

Puede formularse como un teorema de teoría general de los conjuntos, en la forma siguiente:

TEOREMA. *Si* $\langle A, P, E \rangle$ *es una estructura de diferencias finitas iguales, si* $x, y, z \in A$, *y si* $x, y\ E\ x, z$, *entonces* $y = z$.

Con respecto de los teoremas elementales, es una práctica común omitir en la hipótesis la exigencia de que $x, y, z \in A$. Se llega así a la sencilla formulación:

TEOREMA E1′. *Si* $x, y\ E\ x, z$, *entonces* $y = z$.

que es el estilo seguido en los ejercicios; pero debe entenderse que la corrección literal demanda una adición a la hipótesis: que los elementos considerados pertenezcan a A.

Para enunciar el teorema de representación necesitamos definir las estructuras de diferencias numéricas iguales.

DEFINICIÓN 6. *Sea* N *un conjunto finito de números, tal que las diferencias entre números adyacentes, en ordenación natural* $<$, *son iguales. Sea* W *una relación cuaternaria tal que para cualesquiera números* x, y, u *y* v,

$$x, y\ W\ u, v \leftrightarrow x - y = u - v.$$

Entonces $\langle N, <, W \rangle$ *es una estructura de diferencias numéricas iguales finitas.*

TEOREMA 3 (TEOREMA DE LA REPRESENTACIÓN). *Si una estructura de diferencias es una estructura de diferencias iguales*

[15] Las dos principales debilidades de los axiomas, desde el punto de vista de las clases de juicios o respuestas obtenidas de los individuos con motivo de diferencias percibidas o sentidas de alguna característica de estímulos, son las siguientes: (i) Los juicios de orden y de diferencias no son perfectamente transitivos; (ii) en momentos diferentes la misma presentación de estímulos suscitará respuestas diferentes. Son demasiado complicados para discutirlos aquí los métodos para cambiar los axiomas de modo que sean aplicables a cualquiera de estos fenómenos.

finitas, entonces es isomórfica con respecto de una estructura de diferencias numéricas finitas iguales.

En los ejercicios se dan algunas sugestiones relativas a la demostración de este teorema.

Para una teoría de la medición, exigimos no sólo un teorema de representación numérica, sino también un teorema relativo al carácter único (univocidad) de la representación. Para nuestro ejemplo tenemos el siguiente.

TEOREMA 4 (TEOREMA DE LA UNIVOCIDAD). *Dos estructuras de diferencias numéricas finitas iguales cualesquiera, isomórficas con respecto a una estructura dada de diferencias finitas iguales, se relacionan por medio de una transformación lineal. Esto es, si* $\langle N, <, W \rangle$ *y* $\langle N', <, W \rangle$ *son las dos estructuras numéricas, entonces hay un número positivo* α *y un número* β, *tal que para cualquier x en N, hay una x′ única en N′, tal que*

$$x' = \alpha x + \beta.$$ [16]

La segunda aplicación principal en la filosofía de la ciencia de la noción de un teorema de representación es el problema de demostrar que una rama de la ciencia puede ser *reducida* a otra. Por ejemplo, se sabe que gran parte de la termodinámica clásica puede deducirse de la mecánica estadística, en el sentido de que muchas leyes de termodinámica pueden deducirse de varias leyes fundamentales de la mecánica estadística. Para demostrar en un sentido preciso que la termodinámica puede reducirse a la mecánica estadística, necesitaríamos axiomatizar ambas disciplinas, definiendo predicados sinforemáticos adecuados, y a continuación demostrar que, dado un modelo cualquier T de termodinámica, podemos encontrar un modelo de mecánica estadística, que puede tomarse como base para construir un modelo isomórfico con respecto de T. Los ejemplos de esa reducción son demasiado complicados para incluirlos en este texto.[17]

[16] Naturalmente, se obtienen diferentes clases de univocidad para diferentes clases de medición. Por ejemplo, en la teoría de la medición de masas se obtiene univocidad hasta una transformación de similaridad; esto es, desde uno de dos modelos isomórficos hasta un modelo dado de la teoría es posible obtener el otro multiplicando por un número positivo (que intuitivamente corresponde a un cambio en la unidad de medición de la masa). Con respecto a la univocidad, la clase de medición que se obtiene para escalas de diferencias finitas iguales es como la obtenida para la longitud o las mediciones ordinarias de temperatura. Se escogen una unidad y un origen arbitrarios.

[17] Un análisis exacto y detallado de la reducción de la mecánica de los cuerpos rígidos a mecánica de los puntos materiales, se encuentra en Ernest Adams, *Axiomatic Foundations of Rigid Body Mechanics*, disertación, Universidad de Stanford, 1955.

EJERCICIOS

1. Usando la Definición 1, pruébese:

(a) Cualquier álgebra es isomórfica con respecto de sí misma.
(b) Si una álgebra $\mathfrak{A} = \langle A, \circ \rangle$ es isomórfica con respecto de una álgebra $\mathfrak{A}' = \langle A', \circ' \rangle$, entonces $\mathfrak{A}'$ es isomórfica con respecto de $\mathfrak{A}$.
(c) Si una álgebra $\mathfrak{A} = \langle A, \circ \rangle$ es isomórfica con respecto del álgebra $\mathfrak{A}' = \langle A', \circ' \rangle$ y si $\mathfrak{A}'$ es isomórfica con respecto de una álgebra $\mathfrak{A}'' = \langle A'', \circ'' \rangle$, entonces $\mathfrak{A}$ es isomórfica con respecto de $\mathfrak{A}''$.

2. Sea $\mathfrak{A} = \langle A, \circ \rangle$ una álgebra que es un grupo, y sea $\mathfrak{A}$ isomórfica con respecto de una álgebra $\mathfrak{A}' = \langle A, \circ' \rangle$. Demuestre en detalle que $\mathfrak{A}'$ es un grupo.

3. Dése un ejemplo de dos grupos que no sean isomórficos, y demuéstrese que no lo son.

4. Sean $\langle A, R \rangle$ y $\langle B, S \rangle$ simples estructuras de relación. Demuéstrese que son verdaderos cualesquiera de los siguientes enunciados. Dense contraejemplos para los que sean falsos.

(a) $\langle A, \breve{R} \rangle$ es isomórfico con respecto de $\langle A, R \rangle$.
(b) $\langle A \times B, R \times S \rangle$ es isomórfico con respecto de $\langle B \times A, S \times R \rangle$.
(c) Si $R = S$, entonces $\langle A, R \rangle$ es isomórfico con respecto de $\langle B, S/R \rangle$.
(d) Si $R \neq S$, entonces $\langle A, R \rangle$ no es isomórfico con respecto de $\langle B, S \rangle$.
(e) Si $A \subset B$, entonces $\langle A, R \rangle$ no es isomórfico con respecto de $\langle B, S \rangle$.
(f) Si $\langle A, R \rangle$ es isomórfico con respecto de $\langle B, S \rangle$, entonces $\langle A, R/S \rangle$ es isomórfico con respecto de $\langle B, S/R \rangle$.
(g) Si $\langle A, R \rangle$ es isomórfico con respecto de $\langle B, S \rangle$, entonces $\langle A, R/R \rangle$ es isomórfico con respecto de $\langle B, S/S \rangle$.

5. Definimos: *Una álgebra $\mathfrak{A} = \langle A, \circ \rangle$ es totalmente trivial si y sólo si para toda x y y de A, x = y.* Demuéstrese que la teoría de la trivialidad total es categórica. (Recuérdese que si $\langle A, \circ \rangle$ es una álgebra, entonces A es un conjunto no vacío.)

6. Definimos: *Una estructura simple de relación $\langle A, R \rangle$ es una ordenación simple de orden n si y sólo si se satisfacen los axiomas siguientes:*

AXIOMA 1. *El conjunto A tiene exactamente n elementos.*
AXIOMA 2. *R es reflexiva en A.*
AXIOMA 3. *R es antisimétrica en A.*
AXIOMA 4. *R es transitiva en A.*
AXIOMA 5. *R está conectado en A.*

Demuéstrese que es categórica la teoría de las ordenaciones simples de orden n.

7. Sea M un conjunto no vacío. Demuéstrese que el conjunto de todas las transformaciones sobre M es un grupo con respecto de la composición de funciones.

8. Demuéstrense los siguientes teoremas elementales para estructuras de diferencias finitas iguales. Se indican[18] las demostraciones que requieren el

[18] Este principio fundamental puede formularse simbólicamente:

$$\text{(I)} \qquad \varphi(1) \mathbin{\&} (n)(\varphi(n) \rightarrow \varphi(n+1)) \rightarrow (n)\varphi(n).$$

Para demostrar que una fórmula φ es válida para todas las n, necesitamos solamente establecer la hipótesis de (I), a saber, que φ es válida para 1, y que si φ es válida para n, entonces lo es para $n+1$.

uso del principio de inducción matemática. Las inducciones son siempre sobre potencias de J, que pasamos a definir. (Esta definición recursiva es de hecho la apropiada para potencias de cualquier relación binaria.)

$$xJ^1y \leftrightarrow xJy$$

(1) $$xJ^{n+1}y \leftrightarrow (\exists z)(xJ^nz \,\&\, zJy),$$

o en términos de producto relativo:

$$xJ^1y \leftrightarrow xJy$$
$$xJ^{n+1}y \leftrightarrow xJ^n/Jy.$$

Es necesario poner de relieve un aspecto sinforemático. Los teoremas en realidad se refieren a potencias de J restringidas a A, de manera que (1) debe ser reemplazada por:

$$xJ^{n+1}y \leftrightarrow (\exists z)(z \in A \,\&\, xJ^nz \,\&\, zJy).$$

Recuérdese que se entiende tácitamente que todos los elementos mencionados en los teoremas elementales pertenecen a A.

TEOREMA E1. *Si x, y E x, z, entonces $y = z$.* (SUGESTIÓN: Usense particularmente los Axiomas D5, D6 y D8 y dése una demostración indirecta.)

TEOREMA E2. *Si x, x E u, v, entonces $u = v$.*

TEOREMA E3. *Si x, y E u, v, entonces y, x E v, u.*

TEOREMA E4. *Si x, y E u, v & y, z E v, w, entonces x, z E u, w.*

TEOREMA E5. *Si xJ^ny, entonces xPy.* (SUGESTIÓN: Usese la inducción sobre n).[19]

TEOREMA E6. *Si xPy, entonces hay una n (positiva y entera), tal xJ^ny.*

TEOREMA E7. *Si xJ^ny & xJ^nz, entonces $y = z$.* (SUGESTIÓN: Usese la inducción sobre n.)

TEOREMA E8. *Si xJ^my & yJ^nz, entonces $xJ^{m+n}z$.* (SUGESTIÓN: Usese la inducción sobre n.)

TEOREMA E9. *Si xJ^my & $xJ^{m+n}z$, entonces yJ^nz.* (SUGESTIÓN: Usese la inducción sobre n.)

TEOREMA E10. *Si $xJ^{m+n}y$, entonces hay una z (en A), tal que xJ^mz.* (SUGESTIÓN: Usese la inducción sobre m.)

Nótese que los Teoremas de E5 a E10 dependen solamente de los axiomas de D1 a D3.

TEOREMA E11. *Si xJ^ny & uJ^nv, entonces x, y E u, v.* (SUGESTIÓN. Usese la inducción sobre n.)

TEOREMA E12. *x, y E x, y.*

[19] Las demostraciones por inducción que son necesarias para estos teoremas elementales, son todas sencillas en extremo.

Nótese que el Teorema E12 y los Axiomas D5 y D7 juntamente aseveran que E es una relación de equivalencia sobre $A \times A$.

TEOREMA E13. *Si x, y E u, v, entonces o $\exists n$, tal que $xJ^n y$ & $uJ^n v$, o $\exists n$, tal que $yJ^n x$ & $vJ^n u$, o $x = y$ & $u = v$.*

9. Demuéstrese el teorema de representación para estructuras de diferencias finitas iguales (Teorema 3). SUGESTIÓN: Sea z^* el primer elemento de A con respecto de la ordenación P. Defínase la función numérica φ sobre A como sigue, para toda x en A:

$$\varphi(x) = \begin{cases} 1 & \text{si } x = z^* \\ n + 1 & \text{si } z^* J^n x. \end{cases}$$

Usando a continuación los teoremas elementales, demuéstrese:

(i) $xPy \leftrightarrow \varphi(x) < \varphi(y)$
(ii) $x, y \, E \, u, v \leftrightarrow \varphi(x) - \varphi(y) = \varphi(u) - \varphi(v)$.

Demuéstrese (ii) considerando los tres casos listados en el Teorema Elemental E13.

10. Demuéstrese el teorema de univocidad para representaciones numéricas de estructuras de diferencias finitas iguales (Teorema 4). Sean $\mathfrak{N}_1 = \langle N_1, <, W \rangle$ y $\mathfrak{N}_2 = \langle N_2, <, W \rangle$ dos estructuras de diferencias finitas numéricas, isomórficas con respecto de una estructura dada de diferencias finitas iguales $\langle A, P, E \rangle$ y sean φ_1 y φ_2 funciones adecuadas de isomorfismo para $\mathfrak{N}_1$ y $\mathfrak{N}_2$, respectivamente, esto es, $D(\varphi_1) = D(\varphi_2) = A$, $\mathscr{R}(\varphi_1) = N_1$, $\mathscr{R}(\varphi_2) = N_2$, etc. Definimos ahora para toda x de A, dos funciones h_1 y h_2.

$$h_1(x) = \frac{\varphi_1(x) - \varphi_1(z^*)}{\varphi_1(z^{**}) - \varphi_1(z^*)}$$

$$h_2(x) = \frac{\varphi_2(x) - \varphi_2(z^*)}{\varphi_2(z^{**}) - \varphi_2(z^*)},$$

donde z^* es el primer elemento de A bajo la ordenación P y z^{**} el segundo elemento. Demuestre en seguida que (i) h_1 es una transformación lineal de φ_1, h_2 es una transformación lineal de φ_2 y (ii) $h_1 = h_2$. Es entonces fácil demostrar que φ_1 es una transformación lineal de φ_2, esto es, hay unos números α, β, con $\alpha > 0$, tales que para toda x de A

$$\varphi_1(x) = \alpha\varphi_2(x) + \beta.$$

La demostración de la identidad de h_1 y h_2 requiere el uso de la inducción.

§12.4. Ejemplo: Probabilidad. En este apartado deseamos ejemplificar los métodos axiomáticos descritos en los apartados precedentes. Hemos escogido la teoría elemental de los espacios de probabilidad porque el desarrollo de esta teoría exige numerosas aplicaciones simples y directas de relaciones y operaciones de conjuntos. Otros motivos para esta selección son la importancia de la teoría de la probabilidad en todas las áreas de la ciencia empírica y el hecho

histórico de las estrechas conexiones existentes entre los fundamentos de la probabilidad y de la lógica. No es posible discutir dentro de los límites de este apartado los fundamentos generales de la probabilidad, pero se hará el intento de relacionar los desarrollos formales con los problemas fundacionales.

Nuestra axiomatización de la probabilidad procede por definición del predicado sinforemático 'es un espacio de probabilidad finitamente aditivo'. Los axiomas se basan en tres nociones primitivas: un conjunto no vacío de resultados posibles, una familia $\mathfrak{F}$ de subconjuntos de X que representa sucesos o eventos posibles y una función con valores reales P, sobre $\mathfrak{F}$; para $E \in \mathfrak{F}$ $P(E)$ se interpreta como la probabilidad de E. Estas tres nociones pueden ilustrarse por medio de un ejemplo sencillo. Sea X el conjunto de todos los resultados posibles de arrojar dos veces una moneda. Entonces

$$X = \{\langle H, H\rangle, \langle H, T\rangle, \langle T, H\rangle, \langle T, T\rangle\}.$$

Sea $\mathfrak{F}$ la familia de todos los subconjuntos de X. El *evento* de obtener cuando menos una cara es el conjunto

$$A = \{\langle H, H\rangle, \langle H, T\rangle, \langle T, H\rangle\}.$$

El evento de obtener exactamente una cara es el conjunto

$$E = \{\langle H, T\rangle, \langle T, H\rangle\},$$

y así sucesivamente. La idea básica importante es que cualquier evento que podría ocurrir como resultado de arrojar dos veces la moneda, podría ser representado por un subconjunto tal de X, que este subconjunto de X tiene como elementos los resultados posibles, tales que la ocurrencia de cualquiera de ellos implicaría la ocurrencia del evento o suceso. Así, el suceso de obtener cuando menos una cara, ocurre si ambos lanzamientos de la moneda son de hecho caras, si el primero es cara y el segundo cruz, o si el primero es cruz y el segundo cara.

Si la moneda está lúdicamente sana o equilibrada, entonces para cada x de X

$$P(\{x\}) = 1/4.$$

Y la probabilidad de cualquier otro suceso puede obtenerse simplemente sumando el número de elementos de X en el subconjunto. Así, la probabilidad de obtener una cara en dos lanzamientos de la moneda es $P(A)$ y

$$(1)\quad \begin{cases} P(A) = P(\{\langle H, H\rangle, \langle H, T\rangle, \langle T, H\rangle\}) \\ \quad = P(\{\langle H, H\rangle\}) + P(\{\langle H, T\rangle\}) + P(\{\langle T, H\rangle\}) \\ \quad = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}. \end{cases}$$

La propiedad aditiva ejemplificada por (1) es una de las propiedades fundamentales de la probabilidad que postulamos como axioma: la probabilidad de cualquiera de dos sucesos incompatibles es igual a la suma de sus probabilidades individuales, esto es, si $A \cap B = \Lambda$, entonces

$$P(A \cup B) = P(A) + P(B).$$

Puesto que en todo este apartado aplicamos la terminología de la probabilidad a conjuntos que están considerados como sucesos o eventos, será útil contar con una tabla que relacione la notación sinforemática con los términos de la probabilidad. Así, la aserción de que $A \in \mathfrak{F}$ corresponde a aseverar que A es un suceso; la aseveración de que $A \cap B = \Lambda$ a la de que son incompatibles los sucesos A y B. En esta tabla y subsecuentemente usaremos una notación compacta para la intersección o unión de un número finito de conjuntos:

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n$$

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n.$$

Correspondientemente, usaremos después las sumas

$$\sum_{i=1}^{n} P(A_i) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

o, más generalmente, para cualesquiera números a_i, $i = 1, \ldots, n$

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n.$$

También, si X es un conjunto finito y f es una función numérica definida sobre X, entonces

$$\sum_{x \in X} f(x)$$

es justamente la suma de los valores de f para todos los elementos de X. Así, si $X = \{a, b, c\}$

$$\sum_{x \in X} f(x) = f(a) + f(b) + f(c).$$

NOTACION SINFOREMATICA Y TERMINOS DE PROBABILIDAD

TEORÍA DE LOS CONJUNTOS	TEORÍA DE LA PROBABILIDAD[20]
(a) $A \in \mathfrak{F}$	(a) A es un suceso.
(b) $A \cap B = \Lambda$	(b) Los sucesos A y B son incompatibles.
(c) $A_i \cap A_j = \Lambda$ para $i \neq j$, $1 \leq i, j \leq n$	(c) Los sucesos $A_1, A_2, \ldots, A_n$ son incompatibles dos a dos.
(d) $\bigcap_{i=1}^{n} A_i = B$	(d) B es el suceso que ocurre, cuando los sucesos $A_1, A_2, \ldots, A_n$ ocurren juntos.
(e) $\bigcup_{i=1}^{n} A_i = B$	(e) B es el acontecimiento que ocurre cuando menos uno de los acontecimientos $A_1, \ldots, A_n$.
(f) $\sim A$	(f) El suceso ocurre cuando A no ocurre.
(g) $A = \Lambda$	(g) El suceso A es imposible.
(h) $A = X$	(h) El suceso A es inevitable o cierto.
(i) $B \subseteq A$	(i) Si ocurre el suceso B, entonces debe ocurrir A.

En (f) la complementación es, por supuesto, con respecto de X. Así, si A es el acontecimiento o evento consistente en obtener cuando menos una cara, en nuestro ejemplo $\sim A$ es el acontecimiento consistente en no obtener cara y

$$\sim A = X \sim A = \{\langle T, T \rangle\}.$$

La relación expresada por (i) es importante. La inclusión sinforemática de acontecimientos corresponde a la ocurrencia de un acontecimiento que implica la ocurrencia de otro. Así, en nuestro ejemplo, la ocurrencia del acontecimiento de obtener exactamente una cara, implica la del acontecimiento de obtener cuando menos una cara y fácilmente vemos que $E \subseteq A$.

Esparcidas entre los desarrollos formales se encuentran algunas otras observaciones intuitivas que justifican las nociones primitivas particulares escogidas. Nuestro primer axioma sobre espacios

[20] Esta tabla y gran parte del contenido de este apartado están inspirados en el Cap. 1 de la obra clásica de A. N. Kolmogorov, *Fundamentos de la Teoría de la Probabilidad*, edición inglesa, Nueva York, 1950.

de probabilidad será que $\mathfrak{F}$ es un campo de conjuntos sobre X, noción que pasamos a definir.[21]

DEFINICIÓN 1. *$\mathfrak{F}$ es un campo de conjuntos sobre X, si y sólo si $\mathfrak{F}$ es una familia no vacía de subconjuntos de X y para toda A y B de $\mathfrak{F}$*

F1. $A \cup B \in \mathfrak{F}$;
F2. $\sim A \in \mathfrak{F}$.

En otras palabras, un campo de conjuntos de X es una familia no vacía de subconjuntos, ocluida bajo unión y complementación. La prueba del siguiente teorema sobre propiedades elementales de campos y conjuntos se deja como ejercicio.

TEOREMA 1. *Si $\mathfrak{F}$ es un campo de conjuntos sobre X, entonces*

(i) $X \in \mathfrak{F}$,
(ii) $\Lambda \in \mathfrak{F}$,
(iii) Si $A \in \mathfrak{F}$ & $B \in \mathfrak{F}$, entonces $A \cap B \in \mathfrak{F}$,
(iv) Si $A \in \mathfrak{F}$ & $B \in \mathfrak{F}$, entonces $A \sim B \in \mathfrak{F}$.

Como hemos visto por el Teorema 1, los campos de conjuntos están también ocluidos bajo intersección y diferencia de conjuntos. Además, el acontecimiento imposible Λ y el acontecimiento X pertenecen ambos al campo. Debe ser claro que para aplicaciones ordinarias de la teoría de la probabilidad sería difícil trabajar con una familia arbitraria de acontecimientos y se necesita algo como un campo. Por ejemplo, sería extremadamente torpe no estar en posibilidad de tratar la intersección de dos acontecimientos como un acontecimiento, o el complemento de un acontecimiento como un acontecimiento. Cuando X es finito generalmente consideramos el campo que es la familia de todos los subconjuntos de X. Este es, por supuesto, el campo más grande sobre X. El campo más pequeño sobre X es el trivial que tiene como miembro solamente a X y al conjunto vacío.

Definimos ahora el concepto básico de este apartado. Pronto se dará la razón de la frase 'finitamente aditivo'. Nótese que en la definición suponemos la estructura sinforemática de X, $\mathfrak{F}$, y P ya mencionada: X es un conjunto no vacío, $\mathfrak{F}$ es una familia de subconjuntos de X y P es una función sobre $\mathfrak{F}$, valores reales, esto es, una función cuyo alcance es un conjunto de números reales. Así

[21] Decimos que $\mathfrak{F}$ es un campo de conjuntos sobre X y no que $\langle X, \mathfrak{F} \rangle$ es un campo de conjuntos, por deferencia al uso seguido en los escritos técnicos sobre la materia.

pues, la definición se aplica a triples ordenados $\langle X, \mathfrak{F}, P\rangle$ que llamaremos *estructuras de funciones de conjuntos*, puesto que P es una función sobre una familia de conjuntos.

DEFINICIÓN 2. *Una estructura de una función de conjuntos* $\mathfrak{X} = \langle X, \mathfrak{F}, P\rangle$ *es un espacio de probabilidad finitamente aditivo si y sólo si, para toda A y B en* $\mathfrak{F}$

P1. $\mathfrak{F}$ *es un campo de conjuntos sobre* X.
P2. $P(A) \geq 0$.
P3. $P(X) = 1$.
P4. *Si* $A \cap B = \Lambda$, *entonces* $P(A \cup B) = P(A) + P(B)$.

X se llama a menudo el *espacio de la muestra* y la *medida de la probabilidad*. El primer axioma, P1, ya ha sido discutido. El segundo simplemente requiere que la probabilidad de cualquier evento sea un número no negativo y el tercero que la probabilidad del evento cierto sea uno. El axioma final enuncia el supuesto de aditividad ya hecho. Sobre la base de P4 podemos demostrar que la medida de probabilidad P es finitamente aditiva; esto es, podemos demostrar que para $A_1, \ldots, A_n$ eventos incompatibles dos a dos,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Sin embargo, en la teoría general de la probabilidad es deseable, por varias razones fundamentales, tener que si $A_1, A_2, \ldots, A_n$ es una secuencia[22] de eventos incompatibles dos a dos, entonces

$$(2) \qquad P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Una medida de probabilidad que satisface a (2) se dice que es *enumerablemente aditiva*. La Ec. (2) no puede deducirse de la Definición 2; esta es la razón que explica la frase restrictiva 'finitamente aditiva'. Sin embargo, si $\mathfrak{F}$ es finita, (2) es válida. En las aplicaciones aquí consideradas nos restringiremos de hecho a conjuntos X finitos. Después del Teorema 3 siguen algunas observaciones relativas a la aditividad enumerable.

En el teorema siguiente se listan algunas consecuencias elementales de la Definición 2. Puesto que las entidades básicas que se discuten en este apartado son espacios de probabilidad finitamente aditivos, los teoremas y definiciones subsecuentes, relativos a un

[22] Una secuencia es solamente una función cuyo dominio es el conjunto de enteros positivos.

espacio arbitrario de probabilidad, finitamente aditivo, se indican por medio de una letra 'E' y no se hará referencia explícita en estos teoremas y definiciones elementales al espacio básico que se considere. Como se hizo notar en el apartado anterior, esa definición o teorema elemental puede convertirse en un teorema o definición general de teoría de los conjuntos, mediante la adición de la hipótesis: *si* $\langle X, P, A\rangle$ *es un espacio de probabilidad finitamente aditivo* y haciendo ciertos cambios adecuados de menor importancia.

Al formular teoremas elementales usamos cada vez que sea posible el lenguaje de la probabilidad, de preferencia al lenguaje de la teoría de los conjuntos. Por ejemplo, decimos 'para cada evento A', en vez, de 'para cada conjunto A de $\mathfrak{F}$'.

TEOREMA E2.

(i) $P(\Lambda) = 0$.
(ii) *Para todo evento* A, $P(A) + P(\sim A) = 1$.
(iii) *Si los eventos* A_i *son incompatibles dos a dos para* $1 \leq i \leq n$, *entonces*

(1) $$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

DEMOSTRACIÓN. Demostramos únicamente (iii), que requiere el uso de la inducción matemática.[23] Para $n = 1$, está claro que

$$P(A_1) = P(A_1).$$

Nuestra hipótesis inductiva para n es que

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Considérese ahora

$$P\left(\bigcup_{i=1}^{n+1} A_i\right).$$

Puesto que

(2) $$\bigcup_{i=1}^{n+1} A_i = \bigcup_{i=1}^{n} A_i \cup A_{n+1},$$

Tenemos

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^{n} A_i \cup A_{n+1}\right) \qquad \text{Según (2)}$$

[23] Véase la nota No. 18 del Ejercicio 8 del § 12.3 para una formulación del principio de inducción matemática (Pág. 333).

$$= P\left(\bigcup_{i=1}^{n} A_i\right) + P(A_{n+1}) \qquad \text{Según P4 e hipótesis del teorema}$$

$$= \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) \qquad \text{Por hipótesis inductiva}$$

$$= \sum_{i=1}^{n+1} P(A_i) \qquad \text{Q.E.D.}$$

La demostración anterior no es enteramente satisfactoria, puesto que no se ha hecho una discusión previa de identidades sinforemáticas como (2). Sin embargo, intuitivamente está claro que (2) es verdadera. En el presente apartado se supone que son conocidas todas las propiedades elementales que se necesiten de las operaciones aritméticas y con conjuntos.

Definimos ahora la noción de distribución de probabilidad sobre un conjunto finito. La noción puede hacerse extensiva a conjuntos infinitos, bajo ciertas restricciones que no se discutirán aquí.

DEFINICIÓN 3. *Si X es un conjunto finito y no vacío, entonces una distribución de probabilidad sobre X es una función no negativa de valores reales, p, definida sobre X, y tal que*

$$\sum_{x \in X} p(x) = 1.$$

Por ejemplo, considérese el anterior lanzamiento de la moneda, en que $X = \{\langle H, H\rangle, \langle H, T\rangle, \langle T, H\rangle, \langle T, T\rangle\}$. Entonces, para cualquier $x \in X$, $p(x) = 1/4$, suponiendo una moneda equilibrada.

Las distribuciones de probabilidad sobre los conjuntos finitos y sobre los espacios de probabilidad finitamente aditivos están relacionados por medio del teorema siguiente. (Nótese que este teorema no tiene prefijo 'E' antes del número, puesto que no es un teorema sobre un espacio arbitrario de probabilidad, finitamente aditivo.)

TEOREMA 3. *Sea X un conjunto finito, no vacío, sea p una distribución de probabilidad sobre X, y sea $\mathfrak{F}$ una familia de subconjuntos de X, tales que $\mathfrak{F}$ sea un campo de X. Sea P una función de valores reales, definida sobre $\mathfrak{F}$ tal que si $A \in \mathfrak{F}$ entonces $P(A) = \sum_{x \in A} p(x)$. Entonces $\langle X, \mathfrak{F}, P\rangle$ es un espacio de probabilidad finitamente aditivo.*

Dejemos la demostración como ejercicio.

Puesto que las distribuciones de probabilidad parecen ser mucho más sencillas de definir que los espacios de probabilidad finitamente

aditivos, podría pensarse que deberían ser las entidades básicas discutidas en este apartado. Podría prescindirse enteramente de los campos de conjuntos y se obtendría la probabilidad de *cualquier* conjunto sumando las probabilidades de sus elementos individuales. Desgraciadamente esta sencillez se desvanece cuando se consideran conjuntos infinitos, y se suscitan muchos problemas. Está fuera del alcance de este apartado sobre teoría elemental de la probabilidad discutir los métodos matemáticos de hecho usados para el manejo de conjuntos infinitos. Baste con decir que consiste en una natural y relativamente sencilla ampliación de nuestro concepto de espacio de probabilidad finitamente aditivo, a un espacio de probabilidad enumerablemente aditivo, en el sentido de que si cada miembro de una secuencia de conjuntos $\langle A_1, A_2, \ldots \rangle$ está en $\mathfrak{F}$, entonces

$$\bigcup_{i=1} A_i \in \mathfrak{F}.$$

Las operaciones enumerables son necesarias para permitir que queden bien definidas varias importantes operaciones limitadoras.

Pasamos a definir la importante noción de probabilidad condicional.

DEFINICIÓN E4. *Si A y B son eventos y P(A) > 0, entonces*

$$P_A(B) = \frac{P(A \cap B)}{P(A)}.$$

El símbolo '$P_A(B)$' se lee: 'la probabilidad condicional de la ocurrencia del evento B, dada la ocurrencia del evento A', o a menudo, para mayor brevedad, 'la probabilidad condicional de B, dado A'. Una segunda notación para la probabilidad condicional es también útil y usaremos cualquiera de las dos indiferentemente.

DEFINICIÓN E5. *Si A y B son acontecimientos, entonces*

$$P(B|A) = P_A(B).$$

Como ejemplo, consideremos la situación en que una moneda aleatoriamente equilibrada se lanza tres veces.

$$X = \{\langle H, H, H\rangle, \langle H, H, T\rangle, \langle H, T, H\rangle, \langle H, T, T\rangle,$$
$$\langle T, H, H\rangle, \langle T, H, T\rangle, \langle T, T, H\rangle, \langle T, T, T\rangle\}$$

$\mathfrak{F}$ = *conjunto de todos los subconjuntos de X*

$P(\{x\}) = 1/8$ *para toda x de X.*

Sea A el evento de cuando menos una cara en los últimos dos lanzamientos, y sea B el evento de exactamente dos cruces en los tres lanzamientos. Entonces

$$A = \{\langle H, H, H\rangle, \langle H, H, T\rangle, \langle H, T, H\rangle, \langle T, H, H\rangle, \langle T, H, T\rangle, \langle T, T, H\rangle\}$$

$$B = \{\langle H, T, T\rangle, \langle T, H, T\rangle, \langle T, T, H\rangle\}$$

$$A \cap B = \{\langle T, H, T\rangle, \langle T, T, H\rangle\}$$

$$P(A) = 3/4$$

$$P(B) = 3/8$$

$$P(A \cap B) = 1/4,$$

y, en consecuencia,

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/4}{3/4} = 1/3.$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/8} = 2/3.$$

Así, si sabemos que en tres lanzamientos de una moneda de alea limpia se obtuvo cuando menos una cara en los dos últimos lanzamientos, entonces la probabilidad de que ocurran exactamente dos cruces es de $1/3$. La idea intuitiva es simplemente que dos de los seis casos de A están tambien en $A \cap B$. Por otra parte, si sabemos que dos cruces ocurrieron exactamente en los tres lanzamientos, la probabilidad de cuando menos una cara en los últimos dos lanzamientos es de $2/3$. Como lo muestra este ejemplo, en general

$$P(B|A) \neq P(A|B).$$

En el siguiente teorema se resumen las propiedades elementales de las probabilidades condicionales; la demostración se deja como ejercicio.

TEOREMA E4. *Si A, B y C son eventos y $P(A) > 0$, entonces:*

(i) $P_A(B) \geq 0$.
(ii) $P_A(X) = 1$.
(iii) $P_A(A) = 1$.

(iv) Si $B \cap C = \Lambda$, entonces $P_A(B \cup C) = P_A(B) + P_A(C)$.
(v) Si $P(B) > 0$, entonces $P_B(A) = P(A)P_A(B)/P(B)$.
(vi) Si $A \cap B = \Lambda$, entonces $P_A(B) = 0$.
(vii) Si $A \subseteq B$, entonces $P_A(B) = 1$.

Podemos usar las diversas partes del Teorema E4 para demostrar que dado el conocimiento de un evento A con probabilidad positiva, podemos generar un nuevo espacio de probabilidad finitamente aditivo partiendo del espacio dado.

TEOREMA E5. *Si A es un evento y* $P(A) > 0$, *entonces* $\langle X, \mathfrak{F}, P_A\rangle$ *es un espacio de probabilidad finita.*

DEMOSTRACIÓN. Necesitamos demostrar que $\langle X, \mathfrak{F}, P_A\rangle$ satisface los cuatro axiomas de la Definición 2. Ya se ha establecido que $\mathfrak{F}$ es un campo de conjuntos sobre X, puesto que $\langle X, \mathfrak{F}, P\rangle$ es un espacio de probabilidad finitamente aditivo y de este modo se satisface el axioma P1. De la parte (i) del Teorema E4 se sigue que el Axioma P2 queda satisfecho; de la parte (ii) que se satisface el Axioma P3 y de la parte (iv) que se satisface el Axioma P4. Q.E.D.

El siguiente teorema es extremadamente útil en sus aplicaciones; expresa la probabilidad de un evento como una suma de ciertos productos. A menudo se le llama el *Teorema de la Probabilidad Total.*

TEOREMA E6. *Si* $X = \bigcup_{i=1}^{n} A_i$ *y si los eventos* A_i *son incompatibles y* $P(A_i) > 0$, *para* $1 \leq i \leq n$, *entonces, para todo evento B*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

DEMOSTRACIÓN. Usando la hipótesis del teorema y el hecho de que la intersección de los conjuntos es distributiva con respecto de la unión de un número finito de conjuntos, tenemos:

(1) $$B = B \cap X = B \cap \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B \cap A_i.$$

De aquí que

(2) $$P(B) = P\left(\bigcup_{i=1}^{n} B \cap A_i\right).$$

Puesto que por hipótesis $A_i \cap A_j = \Lambda$ para $i \neq j$, podemos aplicar nuestro resultado anterior a la parte de la aditividad finita (iii) del Teorema E2, a (2), y obtener

$$(3) \qquad P(B) = \sum_{i=1}^{n} P(B \cap A_i).$$

Pero por la Definición E5

$$(4) \qquad P(B \cap A_i) = P(B|A_i)P(A_i)$$

y nuestro teorema se sigue inmediatamente de (3) y de (4). Q.E.D.

Un ejemplo de la aplicación del teorema sobre probabilidad total dentro de una teoría científica puede describirse en términos cualitativos considerando la teoría estadística del aprendizaje, tal como ha sido desarrollada en los últimos años por W. K. Estes, R. R. Bush, F. Mosteller y otros. Estamos trabajando con una situación en que un sujeto de experimentación puede dar una de dos respuestas. Por ejemplo, una vuelta a la izquierda o a la derecha, de una rata, en un laberinto *T*; o la predicción hecha por un sujeto humano de que destellará una luz izquierda o derecha. Además, de acuerdo con algún patrón fijado por el experimentador, el sujeto es recompensado por dar ciertas respuestas a ciertas pruebas. El problema general de la teoría es el de predecir qué respuesta dará el sujeto a una prueba cualquiera; la idea general es que un sujeto aprende a dar las respuestas mejor recompensadas. Los axiomas de la teoría precisan esta idea general, enunciando cómo cambian las probabilidades de la respuesta, (i) dada una respuesta recompensada; (ii) dada una respuesta no recompensada, con un sujeto que hace la observación de que la otra respuesta habría sido recompensada; (iii) dada una respuesta recompensada y un sujeto que hace la observación de que ninguna respuesta habría sido recompensada. Hay un axioma para cada uno de los tres casos, exactamente uno de los cuales debe ocurrir en cualquier prueba. Podemos usar los axiomas del aprendizaje y el teorema sobre la probabilidad total para evaluar la probabilidad de una respuesta izquierda, digamos, en la prueba *n*. (Una caracterización más exacta de la teoría estadística del aprendizaje constituiría una digresión demasiado larga.)

A continuación enunciamos y probamos el *teorema de Bayes*, cuyas aplicaciones han sido la fuente de interminables controversias. La prueba es relativamente trivial.

TEOREMA E7. *Si*

$$X = \bigcup_{i=1}^{n} H_i,$$

y si los eventos H_i *son incompatibles dos a dos y* $P(H_i) > 0$, $1 \leq i \leq n$, *y si B es un acontemimiento tal que* $P(B) > 0$, *entonces*:

$$P(H_i|B) = \frac{P(B|H_i)P(H_i)}{\sum_{i=1}^{n} P(B|H_i)P(H_i)}.$$

DEMOSTRACIÓN. Según la parte (v) del Teorema E4,

$$(1) \qquad P(H_i|B) = \frac{P(B|H_i)P(H_i)}{P(B)}.$$

Pero por el Teorema E6

$$(2) \qquad P(B) = \sum_{i=1}^{n} P(B|H_i)P(H_i).$$

La conclusión deseada se sigue inmediatamente de (1) y (2). Q.E.D.

Debe tomarse nota de la terminología generalmente usada en conexión con el teorema de Bayes. Los eventos H_i son llamados *hipótesis*, que son, por supuesto, incompatibles dos a dos y exhaustivos. Una de las hipótesis debería ofrecer una 'mejor explicación' del acontecimiento B. La probabilidad $P(H_i)$ se llama una *probabilidad a priori* de la hipótesis H_i. Estas probabilidades a priori son el centro de la mayor parte de la controversia relativa a aplicaciones del teorema. La probabilidad condicional $P(B\ H_i)$ es llamada la probabilidad de H_i cuando se observa a B.[24] La probabilidad condicional $P(H_i|B)$ se llama la *probabilidad a posteriori* de la hipótesis H_i, dado el evento observado B.

Un ejemplo sencillo ilustrará algunos de los problemas que se suscitan en las aplicaciones de la probabilidad, particularmente en la aplicación del teorema de Bayes. Supóngase que hay una urna que contiene cuatro bolas, de las cuales sabemos que puede ser cierta alguna de las dos cosas siguientes:

H_1: Las cuatro bolas son blancas.
H_2: Dos bolas son blancas y dos son negras.

Supóngase, además, que se extrae al azar una bola y que resulta blanca. Nuestro problema consiste en calcular las probabilidades

[24] El término *probabilidad (inversa)*, común en la estadística moderna, invierte el lenguaje natural de la probabilidad condicional: la probabilidad de B, dado H_i. Las razones de esta inversión se verán más claras en lo que sigue.

a posteriori de H_1 y de H_2. Tres problemas se encuentran en un análisis exacto de la situación:

(i) ¿Cuál es el correcto espacio de muestra X?
(ii) ¿Qué hay que decir sobre las probabilidades a priori de H_1 y de H_2?
(iii) ¿Qué hay que decir sobre las probabilidades (inversas) de H_1 y de H_2?

Pasando al primer problema, vemos inmediatamente que las definiciones y teoremas precedentes de la teoría de la probabilidad no suministran una regla exacta para la construcción del espacio de muestra. Un poco de reflexión indica que puede construirse en una forma tal que cuatro selecciones o tiros consecutivos, sin devolver la bola a la urna, determinarán necesariamente H_1 o H_2, tan ciertamente como que tienen probabilidad 1. Designemos este espacio particular de muestra por 'X_1'. Sabiendo que H_1 o H_2 es cierto, X_1 es simplemente el siguiente conjunto de cuádruples ordenados:

$$X_1 = \{\langle W, W, W, W\rangle, \langle W, W, B, B\rangle, \langle B, B, W, W\rangle, \langle W, B, B, W\rangle, \langle B, W, W, B\rangle, \langle W, B, W, B\rangle, \langle B, W, B, W\rangle\}.$$

Se observará que H_1 y H_2 son los siguientes subconjuntos de X.

$$H_1 = \{\langle W, W, W, W\rangle\}$$

$$H_2 = \{\langle W, W, B, B\rangle, \langle B, B, W, W\rangle, \langle W, B, B, W\rangle, \langle B, W, W, B\rangle, \langle W, B, W, B\rangle\ \langle B, W, B, W\rangle\} = \sim H_1.$$

Si E es el evento de obtener una bola blanca en el primer tiro, entonces

$$E = \{\langle W, W, W, W\rangle, \langle W, B, W, B\rangle, \langle W, W, B, B\rangle, \langle W, B, B, W\rangle\}.$$

El conjunto X_1 no es, sin embargo, el único espacio de muestra correcto para el análisis de nuestro problema de la urna. Está claro que podría decidirse cuál es la hipótesis correcta sobre la base de sólo tres tiros sin devolución. Así, podríamos usar:

$$X_2 = \{\langle W, W, W\rangle, \langle W, W, B\rangle, \langle W, B, W\rangle, \langle B, W, W\rangle, \langle B, B, W\rangle, \langle B, W, B\rangle, \langle W, B, B\rangle\}.$$

Obsérvese que el número de elementos de X_2 es igual al de X_1. Esto es, su cardinalidad es la misma. Si X_2 es el espacio de muestra básico, entonces H_1, H_2 y E son los conjuntos siguientes:

$$H_1 = \{\langle W, W, W\rangle\}$$

$$H_2 = \{\langle W, W, B\rangle, \langle W, B, W\rangle, \langle B, W, W\rangle, \langle B, B, W\rangle, \langle B, W, B\rangle, \langle W, B, B\rangle\} = \sim H_1$$

$$E = \{\langle W, W, W\rangle, \langle W, W, B\rangle, \langle W, B, W\rangle, \langle W, B, B\rangle\}.$$

Por otra parte, si se usaran como elementos del espacio de muestra conjuntos no ordenados, se suscitarían dificultades. Por ejemplo, si usáramos

$$X_3 = \{\{W_1, W_2, W_3, W_4\}, \{W_1, W_2, B_1, B_2\}\}.$$

Entonces

$$H_1 = \{\{W_1, W_2, W_3, W_4\}\}$$

$$H_2 = \{\{W_1, W_2, B_1, B_2\}\},$$

pero no hay un subconjunto apropiado de X_3 para representar el evento de sacar una bola blanca. No obstante, X_1 y X_2 no constituyen las únicas posibilidades legítimas. Por varias razones, muchos estadísticos preferirían el siguiente espacio de muestra.

$$X_4 = \{\langle 4, W\rangle, \langle 2, W\rangle, \langle 2, B\rangle\},$$

y así

$$H_1 = \{\langle 4, W\rangle\}$$

$$H_2 = \{\langle 2, W\rangle, \langle 2, B\rangle\}$$

$$E = \{\langle 4, W\rangle, \langle 2, W\rangle\}.$$

El número que es el primer miembro de cada par ordenado indica el número total de bolas blancas en la urna, en la hipótesis dada (H_1 o H_2). El segundo miembro indica el resultado posible de sacar una bola. Una ventaja de X_4 sobre X_1 y X_2 es que X_4 tiene sólo tres elementos. La reducción del número de elementos es una consecuencia del hecho de que X_4 expresa los resultados posibles de exactamente un tiro, mientras que X_1 indica los posibles resultados de cuatro muestras sin devolución y X_2 el de tres muestras sin devolución. El espacio de muestra X_4 está cortado a la medida de nuestro particular problema de la urna, en una forma en que no están X_1 y X_2.

Tomando X_4 como nuestro espacio de muestra, podemos pasar a nuestro segundo problema, el de determinar las probabilidades a priori $P(H_1)$ y $P(H_2)$. Es una feliz característica de los problemas

de los libros de texto (en oposición a los problemas de la vida real) el que puedan ser tan artificial y tan completamente prescritos. En la descripción de nuestro problema de la urna, no se dijo en absoluto nada sobre H_1 y H_2, a no ser que exactamente una de ellas era cierta. A una situación de esta naturaleza podemos aplicar lo que es, sin duda, la regla más controvertida, de aplicación en el campo de la probabilidad, a saber, el *principio de indiferencia*:[25] Si no hay evidencia para considerar un evento como más probable que otro, asígneseles probabilidades iguales. De este principio resulta que

$$P(H_1) = P(H_2) = \tfrac{1}{2}.$$

Adelante se discute el principio de la probabilidad inversa máxima, estrechamente relacionado al anterior, pero más fácil de aceptar.

Nuestro tercer problema concierne a las probabilidades inversas $P(E|H_1)$ y $P(E|H_2)$. Aunque se necesita una regla de aplicación para evaluarlas, no es ni con mucho tan controvertida como las diversas reglas propuestas para evaluar probabilidades a priori. La regla que tenemos en la mente puede llamarse el *principio de muestreo simple al azar*. Hablando en términos generales, la idea del principio es que usamos un procedimiento de muestreo tal que si H es la hipótesis verdadera, cualquier resultado posible del muestreo, compatible con H, es tan probable como cualquier otro resultado como ese. Así, aquí

$$P(E|H_2) = \tfrac{1}{2}$$

$$P(\sim E|H_2) = \tfrac{1}{2}.$$

Por supuesto, independientemente de cualquier problema de muestreo, en el problema de la urna que se considera,

$$P(E|H_1) = 1.$$

El cálculo de las probabilidades a posteriori $P(H_1|E)$ y $P(H_2|E)$ es ahora directo mediante el uso del teorema de Bayes:

$$\begin{aligned} P(H_1|E) &= \frac{P(E|H_1)P(H_1)}{P(E|H_1)P(H_1) + P(E|H_2)P(H_2)} \\ &= \frac{1\cdot\frac{1}{2}}{1\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}} \\ &= \tfrac{2}{3}, \end{aligned}$$

[25] Generalmente atribuido a James Bernouilli, *Ars Conjectandi*, Basel, 1713; pero su prominencia en la historia de la probabilidad se debe más al uso del mismo hecho por

y de modo semejante encontramos que

$$P(H_2|E) = \tfrac{1}{3}.$$

Obviamente cualquiera que use el principio de indiferencia favorecería la hipótesis H_1, sobre la base de la muestra E.

Es importante advertir que el uso del principio del muestreo al azar es sensible al espacio de muestra particular adoptado para el análisis de un problema. Por ejemplo, para cambiar ligeramente nuestro problema, sea H_3 la hipótesis de que tres de las cuatro bolas de la urna son blancas y la cuarta es negra. Supóngase, además, por el momento, que H_1 y H_3 son hipótesis exhaustivas. Tendríamos un espacio de muestra análogo a X_4

$$X_5 = \{\langle 4, W\rangle, \langle 3, W\rangle, \langle 3, B\rangle\}$$

y

$$H_1 = \{\langle 4, W\rangle\}$$

$$H_3 = \{\langle 3, W\rangle, \langle 3, B\rangle\}$$

$$E = \{\langle 4, W\rangle, \langle 3, W\rangle\},$$

pero, entonces, tendríamos el siguiente caso de probabilidad inversa:

$$P(E|H_3) = \tfrac{3}{4},$$

puesto que H_3 es verdadero, de las cuatro bolas tres son blancas y de este modo la probabilidad de extraer una bola blanca es de $\tfrac{3}{4}$.

El carácter insatisfactorio de los métodos relativamente vagos para aplicar el principio de muestreo al azar puede ser considerablemente reducido usando un espacio de muestra de estructura semejante a X_4 y X_5, pero más concreto. La idea es la de asignar nombres a cada una de las bolas (como lo hicimos tácitamente para el espacio insatisfactorio X_3). Si H_1 es verdadero, sean W_1, W_2, W_3, W_4 las cuatro bolas. Si H_3 es verdadero, sean W_1, W_2, W_3, B_1 las cuatro bolas. Nuestro espacio de muestra es entonces:

$$X_6 = \{\langle 4, W_1\rangle, \langle 4, W_2\rangle, \langle 4, W_3\rangle, \langle 4, W_4\rangle, \langle 3, W_1\rangle, \langle 3, W_2\rangle, \langle 3, W_3\rangle, \langle 3, B_1\rangle\},$$

y es obvio que subconjuntos de X_6 son H_1, H_3 y E. Por *evento atómico*

Pierre Laplace en su gran tratado *Théorie analytique des probabilités*, París, 1812; 2a. Ed., 1814. El principio de indiferencia es llamado también el principio de razón no suficente. No es posible discutir aquí la relativa legitimidad del principio, ni analizar algunas de las paradojas a que ha dado origen.

entendemos un subconjunto unidad de X. Así, $\{\langle 3, W_2\rangle\}$ es un evento atómico, por ejemplo. El principio de muestreo al azar puede ahora expresarse muy sencillamente. Supóngase que hay n eventos atómicos incluidos en H. Entonces la probabilidad inversa dada H de un evento atómico A incluido en H es justamente:

$$P(A|H) = \frac{1}{n}.$$

Y para X_6 tenemos entonces sobre esta base:

$$\begin{aligned} P(E|H_3) &= \sum_{i=1}^{4} P(\{\langle 4, W_i\rangle\}|H_3) + \sum_{i=1}^{3} P(\{\langle 3, W_i\rangle\}|H_3) \\ &= 0 + 1/4 + 1/4 + 1/4 \\ &= 3/4. \end{aligned}$$

Será de cierto interés considerar un problema de una urna, un poco más complicado, que analizamos sin tantas digresiones. Supóngase de nuevo que la urna contiene cuatro bolas y que H_2 o H_3 es verdadero, esto es, que la urna contiene dos bolas blancas y dos bolas negras o tres bolas blancas y una negra. Supóngase, además, que se extrae al azar una muestra, consistente en una bola, sin devolverla a la urna, y a continuación dos bolas con devolución y que se observa que la primera bola es blanca y las dos siguientes negras. Podríamos tomar como elementos de nuestro espacio de muestra cuádruples ordenados de la forma $\langle 2, W_1, B_1, B_2\rangle$, pero puesto que, conforme al principio de muestreo al azar que suponemos, no importa el orden de las dos bolas con devolución, en la doble muestra, podemos reducir el número de elementos del espacio de muestra usando triples ordenados de la forma $\langle 2, W_1, \{B_1, B_2\}\rangle$, donde el último miembro de cada triple es un conjunto, a saber, el conjunto de bolas extraídas en la muestra doble con devolución. Considerando todas las combinaciones posibles de resultados para H_2 y H_3, vemos que el espacio de muestra consta de cuarenta y ocho elementos:

$$\begin{aligned} X_7 = \{&\langle 2, W_1, \{W_2, B_1\}\rangle, \langle 2, W_1, \{W_2, B_2\}\rangle, \langle 2, W_1, \{B_1, B_2\}\rangle, \\ &\langle 2, W_1, \{W_2\}\rangle, \langle 2, W_1, \{B_1\}\rangle, \langle 2, W_1, \{B_2\}\rangle, \ldots, \\ &\langle 2, B_2, \{B_1\}\rangle, \langle 3, W_1, \{W_2, W_3\}\rangle, \ldots, \langle 3, B_1, \{W_3\}\rangle\}. \end{aligned}$$

Nótese que un elemento como $\langle 2, W_1, \{W_2\}\rangle$ representa el hecho de que la bola W_2 sea extraída dos veces; pues

$$\{W_2, W_2\} = \{W_2\},$$

usamos el conjunto unidad.

H_2 es simplemente el conjunto de elementos cuyos primeros miembros son 2, y H_3 aquellos cuyos primeros elementos son 3. La muestra F consistente en una bola blanca sin devolución y a continuación dos bolas negras con devolución es el conjunto.

$$\begin{aligned} F = \{&\langle 2, W_1, \{B_1, B_2\}\rangle, \langle 2, W_1, \{B_1\}\rangle, \langle 2, W_1, \{B_2\}\rangle, \\ &\langle 2, W_2, \{B_1, B_2\}\rangle, \langle 2, W_2, \{B_1\}\rangle, \langle 2, W_2, \{B_2\}\rangle, \\ &\langle 3, W_1, \{B_1\}\rangle, \langle 3, W_2, \{B_1\}\rangle, \langle 3, W_3, \{B_1\}\rangle\}. \end{aligned}$$

Suponiendo de nuevo el principio de indiferencia,

$$P(H_2) = P(H_3) = 1/2.$$

Además, usando el principio de muestreo al azar, fácilmente computamos las probabilidades inversas.

$$P(F|H_2) = 6/24 = 1/4,$$

$$P(F|H_3) = 3/24 = 1/8.$$

Entonces tenemos para las probabilidades a posteriori:

$$\begin{aligned} P(H_2|F) &= \frac{P(F|H_2)P(H_2)}{P(F|H_2)P(H_2) + P(F|H_3)P(H_3)} \\ &= \frac{1/4 \cdot 1/2}{1/4 \cdot 1/2 + 1/8 \cdot 1/2} \\ &= 2/3, \\ P(H_3|F) &= 1 - P(H_2|F) = 1/3. \end{aligned}$$

Suponiendo el principio de indiferencia, la evidencia F claramente favorece la hipótesis H_2.

No ha sido posible presentar aquí una crítica del principio de indiferencia, pero podemos mencionar el *principio de máxima probabilidad inversa*, de R. A. Fisher, que elude la necesidad de establecer probabilidades a priori. En el caso de nuestros problemas de urnas, el principio dice que se debe aceptar la hipótesis cuya probabilidad inversa es máxima, esto es, máxima para el ejemplo

observado. Para los problemas de urnas considerados resulta también inmediatamente evidente qué hipotesis tiene la probabilidad inversa máxima. La aplicación del principio a la estimación de alguna cantidad, como la cantidad de peces en un lago, sobre la base de una muestra, es más complicada.

Al discutir estos sencillos problemas de urnas, se ha hecho hincapié en la construcción de un espacio de muestra apropiado y se ha pasado por alto la selección de un campo de conjuntos, pues se ha supuesto tácitamente que el campo de conjuntos es el conjunto de todos los subconjuntos del espacio de muestra, selección que casi siempre es apropiada cuando el espacio de muestra tiene un número finito de elementos. De modo semejante, la medida de probabilidad P no ha sido completamente especificada, sino solamente para los argumentos importantes para el problema tratado. En los problemas considerados está claro que el principio de indiferencia y el principio de muestreo al azar determinan completamente a P, pero a menudo no es éste el caso. Por ejemplo, si se usa el principio de máxima probabilidad inversa y se rechaza el principio de indiferencia, no es posible determinar completamente a P con los datos suministrados, en los anteriores problemas de urnas. Por esta razón, aun cuando sea fijo el espacio de muestra X, hay frecuentemente muchos espacios de probabilidad finitamente aditivos, adecuados para el análisis de un problema dado. Esta situación no es privativa de la teoría de la probabilidad, sino que en general se presenta en la aplicación de cualquier teoría exactamente formulada, a problemas descritos en forma intuitiva e informal.

En este apartado nos ha sido posible considerar solamente un fragmento de la teoría de la probabilidad, que es uno de los temas más ricos y profundos de la matemática moderna.[26] Igualmente, es lamentable, desde un punto de vista filosófico, la omisión de una discusión detallada de los fundamentos de la probabilidad, que ha sido uno de los tópicos más importantes de la filosofía sistemática desde el siglo dieciocho. No podemos discutir aquí ninguna de las dos escuelas de pensamiento: los que sostienen el punto de vista objetivo de que las probabilidades de los eventos se miden por las frecuencias relativas de sus ocurrencias, y los que sostienen el punto de vista subjetivo de que esas probabilidades son primera y principalmente medidas de la creencia razonable de que ocurrirán los acontecimientos.[27]

[26] Los desarrollos comenzados en este capítulo son llevados lejos por W. Feller, en su *An Introduction to Probability Theory and Its Applications*, Vol. I, Nueva York, 1950. De este libro puede pasar el lector a tratados más avanzados sobre el tema de la probabilidad.

[27] Para el punto de vista de la frecuencia relativa véase la obra de von Mises, *Probability, Statistics and Truth*, traducción inglesa, Nueva York, 1939; E. Nagel, *Principles of the*

EJERCICIOS

1. Demuéstrese el Teorema 1.
2. Formúlese de nuevo la Definición 2 en el estilo de la Definición *A* del Apartado 2.
3. Demuéstrese (i) y (ii) del Teorema E2.
4. Demuéstrese el Teorema 3.
5. Demuéstrese el Teorema E4.
6. Considérese la siguiente situación. Tenemos una urna que contiene cinco bolas. Sabemos que una de las dos hipótesis siguientes es verdadera:

   H_1: dos bolas son blancas y tres son negras.
   H_2: tres bolas son blancas y dos son negras.

Para cada uno de los siguientes supuestos construya un espacio de muestra y, suponiendo los principios de indiferencia y muestreo al azar, compute las probabilidades a posteriori de las dos hipótesis.

(a) Extraemos al azar una muestra con devolución, consistente en dos bolas blancas y una negra.
(b) Extraemos al azar una muestra del siguiente tipo: una bola negra sin devolución y a continuación una bola blanca.
(c) Extraemos una muestra al azar del siguiente tipo: Una muestra de dos sin devolución, consistente en una bola negra y una blanca, y a continuación una sola muestra, consistente en una bola negra.

7. Dada la situación del Ejercicio 6, y el permiso para extraer una muestra al azar de dos bolas, ¿qué método de muestreo recomendaría: extraer la primera bola con devolución, o extraerla sin devolución? Justifique su respuesta.

8. Suponiendo que valga $\frac{1}{2}$ la probabilidad de que un niño sea varón, construya el espacio de probabilidad finitamente aditivo apropiado para una familia de cinco niños. ¿Cuál es la probabilidad de que

(a) todos los niños sean varones?
(b) haya dos niños y tres niñas?
(c) haya cuando menos un niño?
(d) haya cuando menos una niña, dado que haya cuando menos tres niños?
(e) haya exactamente dos niños, dado que haya cuando menos dos niñas?

9. Un problema clásico de inducción es el siguiente. Un hombre extrae al azar m bolas con devolución de una urna que contiene N bolas blancas y negras y es posible que no haya bolas blancas o no haya negras. Supóngase que se extraen m bolas negras. ¿Cuál es la probabilidad de que en el tiro $(m+1)$ se extraiga una bola negra?

Theory of Probability, International Encyclopedia of Unified Science, Vol. I, No. 6, Chicago, 1939; H. Reichenbach, *The Theory of Probability*, traducción inglesa, Los Angeles, 1939. Para el punto de vista subjetivo, véase H. Jeffreys, *Theory of Probability*, 2a. Ed., Oxford, 1948; L. J. Savage, *Foundation Statistics*, Nueva York, 1954. Para una variante de la teoría subjetiva estrechamente conectada con la lógica, véase R. Carnap, *Logical Foundations of Probability*, Chicago, 1950.

(a) Sea $N = 4$ y $m = 2$, y constrúyase el espacio de probabilidad finitamente aditivo adecuado, suponiendo los principios de indiferencia y muestreo al azar. Nótese que hay cinco hipótesis mutuamente exclusivas y exhaustivas, puesto que hay cinco posibilidades (de cero a cuatro) para el número de bolas negras.

(b) Demuéstrese que la respuesta general es:

$$\frac{\sum_{i=0}^{N}\left(\frac{i}{N}\right)^{m+1}}{\sum_{i=0}^{N}\left(\frac{i}{N}\right)^{m}}.$$

[Puede demostrarse que a medida que N aumenta sin límite, la probabilidad de una bola negra sobre el tiro $(m+1)$ llega a ser $(m+1)/(m+2)$.]

§12.5. Ejemplo: mecánica. Una de las más exactas y bien establecidas ramas de la ciencia empírica es la teoría clásica de la mecánica de los puntos materiales. En realidad, durante un periodo de casi doscientos años, desde el comienzo del siglo dieciocho hasta el fin del siglo diecinueve se sostuvo ampliamente que la mecánic de los puntos materiales era la teoría fundamental del universo. En este apartado deseamos estudiar brevemente una axiomatización de esta teoría con el fin de demostrar que hasta una teoría relativamente complicada puede recibir una clara formulación dentro de la teoría de los conjuntos. Desgraciadamente, nuestro desarrollo de la mecánica debe usar ciertos conceptos matemáticos que no resultaría cómodo explicar aquí *desde sus principios*.[28] En este momento definiremos las más importantes nociones necesitadas. Un *intervalo* de números reales es el conjunto de números reales existentes entre dos números, mayores que un número dado o menores que otro número dado. Más exactamente, si a y b son números reales $[a, b]$ es el conjunto de todos los números reales x, tales que $a \leq x \leq b$. El intervalo $[a, b]$ es *cerrado*, puesto que incluye sus puntos extremos a y b. Para designar intervalos cerrados se usan, generalmente, corchetes, y paréntesis para designar intervalos *abiertos*. Así, (a, b) es el intervalo abierto que es el conjunto de todos los números x tales que $a < x < b$. Como era de esperarse $(a, b]$ es un intervalo semiabierto, semicerrado, etc. $(-\infty, a)$ es el conjunto de todos los números reales x, tales que $x < a$, y el intervalo $[a, \infty)$ es el conjunto de todos los números reales x tales que $x \geq a$. Finalmente, el intervalo $(-\infty, \infty)$ es el conjunto de todos los números reales.

Un *vector* (*tridimensional*) es un triple ordenado de números reales. Así, $\langle 4, 7, 3\rangle$ es un vector. La interpretación física usual de los vectores es que representan la magnitud y dirección de alguna

[28] Este es el único apartado con respecto del cual el presente libro no es autosuficente.

cantidad física, como la velocidad y la fuerza, con respecto de tres ejes espaciales, mutuamente perpendiculares. Las operaciones binarias con vectores, o con vectores y números reales, se definen en forma natural. Así, si $x = \langle x_1, x_2, x_3 \rangle$ es un vector y a es un número real,

$$ax = a\langle x_1, x_2, x_3\rangle = \langle ax_1, ax_2, ax_3\rangle = xa.$$

Si $x = \langle x_1, x_2, x_3 \rangle$ y $y = \langle y_1, y_2, y_3 \rangle$ son vectores, entonces

$$x + y = \langle x_1, x_2, x_3\rangle + \langle y_1, y_2, y_3\rangle = \langle x_1 + y_1, x_2 + y_2, x_3 + y_3\rangle$$

y

$$x - y = \langle x_1, x_2, x_3\rangle - \langle y_1, y_2, y_3\rangle = \langle x_1 - y_1, x_2 - y_2, x_3 - y_3\rangle.$$

Además,

$$-x = -\langle x_1, x_2, x_3\rangle = \langle -x_1, -x_2, -x_3\rangle.$$

La operación de *producto interno* o *escalar* de los vectores x y y, da un número real y se designa por medio de un punto. Así

$$x \cdot y = \langle x_1, x_2, x_3\rangle \cdot \langle y_1, y_2, y_3\rangle = x_1y_1 + x_2y_2 + x_3y_3.$$

La operación de *producto vectorial* o *de cruz* con los vectores x y y es un vector y se designa con '×' (que no debe confundirse con la operación de producto cartesiano de los conjuntos). Así

$$\begin{aligned} x \times y &= \langle x_1, x_2, x_3\rangle \times \langle y_1, y_2, y_3\rangle \\ &= \langle x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1\rangle. \end{aligned}$$

Si $x = \langle x_1, x_2, x_3 \rangle$ es un vector, entonces

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

El número $|x|$ es el *valor absoluto* del vector x. Las propiedades formales de estas diversas operaciones se enuncian en los ejercicios.

Pueden ser útiles algunas observaciones intuitivas sobre las últimas tres operaciones. Cuando, por ejemplo, x es un vector de velocidad, entonces $|x|$ es la magnitud (independientemente de la dirección y el sentido). El producto escalar tiene la siguiente interpretación geométrica:

$$x \cdot y = |x|\,|y| \cos (x, y),$$

donde $\cos(x, y)$ es el coseno del ángulo que forman entre sí los dos vectores. El valor del producto vectorial es dado por la ecuación:

$$|x \times y| = |x|\,|y|\,|\mathrm{sen}(x, y)|.$$

La dirección del producto vectorial es perpendicular al plano formado por los vectores x y y. Si f es una función con valores reales o vectoriales, entonces Df es la *derivada* de f. No estarán de más unas cuantas observaciones sobre derivadas. Sea f una función que se define en un intervalo T de números reales y cuyo alcance es un conjunto de números reales. Entonces Df es una función cuyo dominio de definición es un subconjunto de T. Si $a \in T$, entonces a está en el dominio de Df si y sólo si hay un número l tal que para toda $\epsilon > 0$, hay una δ tal que si $x \in T$, $x \neq a$ y $|x - a| < \delta$, entonces

$$\left| \frac{f(x) - f(a)}{x - a} - l \right| < \epsilon.$$ [29]

Además, cuando a está en el dominio de Df, se dice que la función f es diferenciable en a. Si f es diferenciable en cada punto de T,[30] se dice que es *diferenciable sobre o a lo largo de T*. Hemos introducido explícitamente estas definiciones para evitar frases usuales como 'la derivada de f existe en a'. Aunque este lenguaje de existencia es perfectamente claro para el lector experimentado, no es una expresión feliz desde un punto de vista lógico y, en consecuencia, la evitaremos. La derivada segunda de una función f es simplemente la función $D(Df)$, que designaremos como D^2f. Si el intervalo T es el dominio de definición de una función f, entonces f es *doblemente diferenciable* en el punto a de T, si a está en el dominio de D^2f, etc. Uno de nuestros axiomas de mecánica es que la función de posición de una partícula es doblemente diferenciable en el intervalo de tiempo del sistema mecánico.

Pasamos ahora a nuestras seis nociones primitivas de la mecánica de puntos materiales. P y T son conjuntos, s es una función binaria, m es una función unaria, f es una función ternaria y g es una función binaria.

La interpretación física que se intenta hacer de P es que es el conjunto de *partículas*. T se interpreta físicamente como un conjunto de números reales que miden tiempos pasados. Si $p \in P$ y $t \in T$, entonces $s(p, t)$ es un vector que se interpreta físicamente como la *posición* del punto material p en el instante t. Para cada p

[29] Esta definición se extiende fácilmente a funciones de valores vectoriales. Si a está en el dominante de Df, entonces $Df(a) = l$.

[30] En estos contextos, 'punto' se usa a menudo como sinónimo de 'número'.

de P será conveniente introducir la función de posición s_p.[31] Nótese que la s primitiva fija la selección de un sistema de coordenadas. Estas primeras tres nociones primitivas son las nociones cinemáticas de la mecánica de puntos materiales.

Si $p \in P$, entonces $m(p)$ se interpreta como el valor numérico de la *masa* de la partícula p. Si $p, q \in P$, y $t \in T$, entonces $f(p, q, t)$ es la *fuerza* que la partícula q ejerce sobre el punto material p en el instante t. Esas fuerzas se llaman las fuerzas *internas* del sistema mecánico. Si $p \in P$ y $t \in T$, entonces $g(p, t)$ es la *fuerza externa* resultante que actúa sobre la partícula p en el instante t. Por ejemplo, considérense los dos sistemas de puntos materiales consistentes en los planetas Tierra y Venus. Entonces la fuerza interna ejercida sobre la Tierra es la atracción gravitacional de Venus y la fuerza externa resultante sobre la Tierra es la suma vectorial de las fuerzas atractivas gravitacionales de todos los demás cuerpos del sistema solar (y para el caso, del resto del universo). Hay buenas razones para insistir en que las fuerzas externas *dadas* que actúan sobre un punto material sean primitivas y no la fuerza resultante, mayormente si se considera que esta última es definible en términos de la anterior, pero no recíprocamente. Sin embargo, es formalmente más sencillo tratar sólo la fuerza externa resultante sobre un punto material.[32] Las nociones de masa, fuerza interna, y fuerza externa son las nociones *dinámicas* primitivas de la mecánica de puntos materiales.

Estamos ahora listos para enunciar nuestros axiomas cinemáticos y dinámicos, para sistemas de mecánica de puntos materiales.

DEFINICIÓN 1. *Un sistema* $\mathfrak{P} = \langle P, T, s, m, f, g\rangle$ *es un sistema de mecánica de partículas si y sólo si se satisfacen los siguientes siete axiomas*:

AXIOMAS DE CINEMATICA

AXIOMA P1. *El conjunto P es finito y no vacío.*

AXIOMA P2. *El conjunto T es un intervalo de números reales.*

---

[31] Usando la notación lambda del § 11.3, podríamos escribir: $(\lambda t)s(p, t)$, en vez de: s_p.

[32] Para una discusión de este y otros puntos, véase J. C. C. McKinsey, A. C. Sugar y Patrick Suppes, "Axiomatic foundations of classical particle mechanics", *Journal of Rational Mechanics and Analysis*, Vol. 2 (1953), Págs. 253 a 272. La axiomatización dada en este trabajo está estrechamente relacionada con la que se da aquí. Para una axiomatización de la mecánica relativista de puntos materiales véase H. Rubin y P. Suppes, "Transformations of systems of relativistic particle mechanics", *Pacific Journal of Mathematics*, Vol. 4 (1954), Págs. 563 a 601.

AXIOMA P3. *Para p de P, s_p es dos veces diferenciable sobre T.*

AXIOMAS DE DINAMICA

AXIOMA P4. *Para p de P, m(p) es un número positivo real.*

AXIOMA P5. *Para p y q de P y t de T,*

$$f(p, q, t) = -f(q, p, t).$$

AXIOMA P6. *Para p y q de P y t de T,*

$$s(p, t) \times f(p, q, t) = -s(q, t) \times f(q, p, t).$$

AXIOMA P7. *Para p de P y t de T,*

$$m(p)D^2s_p(t) = \sum_{q\in P} f(p, q, t) + g(p, t).$$

Aunque no es posible efectuar aquí una discusión completa de cada axioma y de su interpretación física, pueden hacerse algunas observaciones que aclaran las cosas.

La condición contenida en el primer axioma de que P sea finita, es necesaria para garantizar que estén bien definidas la masa y la energía cinética de todo el sistema. Puesto que no hay interés en que P esté vacío, hemos seguido la tradición del álgebra y hecho a P no vacío. (Recuérdese que en el § 12.2 hemos exigido que para que $\langle A, \circ \rangle$ sea un álgebra, A sea no vacío).

Con respecto del segundo axioma, podría estimarse como más sencillo tomar el intervalo de tiempo como el conjunto de los números reales, pero esto no es conveniente en numerosas aplicaciones, porque, como diría Aristóteles, los puntos materiales continuamente están llegando a ser y desapareciendo. Además, nuestros axiomas no son adecuados para tratar impactos, pero podemos tratar la noción el punto material anterior al instante del impacto, seleccionando el intervalo de tiempo apropiado.

La doctrina filosófica del convencionalismo,[33] esto es, la doctrina de que ciertos supuestos de la ciencia empírica se hacen sólo por motivos de comodidad o conveniencia, es ilustrada por el segundo axioma. Desde el punto de vista de la susceptibilidad de comprobación empírica, este axioma podría ser reemplazado por la exigencia de que T sea un conjunto de números racionales en algún intervalo. Ciertamente es imposible construir un reloj para el cual la medición

[33] El famoso matemático y filósofo francés Henri Poincaré (1854-1912) está prominentemente asociado con esta doctrina.

de tiempos transcurridos no pueda ser representada satisfactoriamente por números racionales. Por otra parte, el supuesto de que T es un conjunto de números racionales es matemáticamente inconveniente, pues cierra el paso a la aplicación directa de los métodos tipo del análisis matemático, en particular a los métodos del cálculo diferencial e integral.

El requerimiento del tercer axioma de que la función de posición de una partícula sea dos veces diferenciable con respecto de T es también una idealización, pues cualquier conjunto de datos empíricos a lo largo de la trayectoria de un punto material pueden ser aproximados al arbitrio por una función que en ningún punto de T es dos veces diferenciable. Una función de esta última clase es completamente imposible de manejar matemáticamente; si esa función se usara para representar el movimiento de una partícula, muchos de los más sencillos problemas de mecánica se volverían extremadamente difíciles, si no es que imposible de resolver. Por otra parte, hay condiciones susceptibles de desarrollo que pueden utilizarse, más débiles que la de doble diferenciabilidad.[34] Hemos caracterizado los primeros tres axiomas como los axiomas cinemáticas de la mecánica. Maxwell da la siguiente sucinta caracterización de cinemática.[35]

> Comenzamos con la cinemática o la ciencia del puro movimiento. En esta división de la disciplina, las ideas que se nos presentan son las de espacio y tiempo. El único atributo de la materia que surge ante nosotros es el de su continuidad de existencia en el espacio y en el tiempo, a saber, el hecho de que toda partícula de materia, en cualquier instante de tiempo, está en un lugar y sólo en uno y que su cambio de lugar durante cualquier intervalo de tiempo lo efectúa moviéndose a lo largo de una trayectoria continua.
>
> Ni la fuerza que afecta el movimiento del cuerpo ni la masa del mismo, de la cual depende la cantidad de fuerza requerida para efectuar el movimiento, se toman en cuenta en la ciencia del movimiento puro.

La exigencia de Maxwell de una trayectoria continua no es suficientemente fuerte, pues hay funciones que son continuas en todos sus puntos, pero en ninguno de ellos son diferenciables. Si ese tipo de función se usara para representar el movimiento de una partícula, la velocidad instantánea de ésta no quedaría definida en ningún punto, puesto que la función velocidad de una partícula o punto material es la primera derivada de su función de posición.

El Axioma P4 requiere que la masa de una partícula sea siempre positiva. Si la masa de algunas partículas fuera igual a cero, entonces,

[34] Véase, por ejemplo, G. Hamel, "Die Axiome der Mechanik", *Handbuch der Physik*, Vol. 5, Págs. 1-42. Sin embargo, la axiomatización de Hamel no satisface el criterio del conjunto teórico desarrollado en este capítulo.

[35] J. Clerk Maxwell, *Matter and Motion*, edición de 1920, Londres, Pág. 79.

por aplicación del axioma final, P7, la aceleración de la partícula no podría ser determinada a partir del conocimiento de las fuerzas que actúan sobre la misma. En la teoría de los cohetes, por ejemplo, es deseable permitir que la masa varíe con el tiempo (lo que corresponde al consumo de combustible), pero haciendo a la masa independiente del tiempo hemos permanecido dentro del marco de la mecánica clásica.

Los Axiomas P5 y P6 suministran una exacta formulación de la tercera ley del movimiento de Newton:[36] a cada acción corresponde siempre una reacción igual y opuesta; o, las reacciones mutuas de dos cuerpos son siempre iguales y de sentido contrario. El quinto axioma corresponde a lo que Hamel llama el *primer principio de la reacción completa* y el sexto axioma al *segundo principio de la reacción completa*.[37] El contenido intuitivo de los dos axiomas debe resultar claro. El quinto axioma requiere que la fuerza ejercida por q sobre p sea igual y opuesta a la ejercida por p sobre q. El sexto axioma requiere que la dirección de las dos fuerzas internas sea a lo largo de la línea que conecta la posición de las dos partículas, que esto se expresa por la ecuación del producto vectorial de P6 puede verse por las siguientes consideraciones. Dos vectores son paralelos si y sólo si hay un número real a tal que

$$ax = y.$$

(Nótese que el vector $\langle 0, 0, 0\rangle$ es paralelo a cualquier vector).[38] Consecuentemente, cuando dos vectores x y y son paralelos, el producto vectorial es el vector cero $\langle 0, 0, 0\rangle$, para

$$\langle ax_1, ax_2, ax_3\rangle = \langle y_1, y_2, y_3\rangle,$$

esto es,

$$ax_i = y_i \qquad \text{para } i = 1, 2, 3,$$

de donde

$$x_2y_3 - x_3y_2 = ax_2x_3 - ax_3x_2 = 0,$$

[36] Isaac Newton, *Principia*, traducción de Cajori, Berkeley, Calif., 1934, Pág. 13. Los *Principia* fueron publicados por primera vez en 1687. Después de los *Elementos* de Euclides ha sido probablemente el trabajo más influyente en la historia de la ciencia. Con respecto de la tercera ley, Newton hace la observación de que "Cualquier cosa que ejerce una tracción o una presión sobre otra es a su vez sometida a tracción o a presión, con la misma intensidad, por la otra. Si se comprime una piedra con el dedo, éste también es comprimido por la piedra. Si un caballo tira de una piedra atada a una cuerda, el caballo (si puedo decirlo), será también tirado hacia la piedra, pues la cuerda tensa, en virtud del esfuerzo por distenderse, tirará del caballo hacia la piedra, con la misma intensidad con que ejerce tracción hacia el caballo, y obstruirá el progreso de lo uno en la misma medida en que promueva el de lo otro."

[37] Hamel, obra citada, Pág. 25.

[38] Algunos autores definen el paralelismo sólo para vectores no cero.

y de modo semejante para los otros dos componentes del producto vectorial. Ahora bien, el axioma enuncia que

$$(1) \qquad s(p, t) \times f(p, q, t) = -s(q, t) \times f(q, p, t).$$

Sumamos a cada miembro de la Ec. (1) $- s(q, t) \times f(p, q, t)$ y usamos el hecho de que el producto vectorial es distributivo con respecto de la adición para obtener

$$(2) \qquad (s(p, t) - s(q, t)) \times f(p, q, t) = -s(q, t) \times (f(q, p, t) + f(p, q, t)).$$

Según el Axioma P5

$$f(q, p, t) + f(p, q, t) = \langle 0, 0, 0 \rangle,$$

de donde inferimos de (2):

$$(s(p, t) - s(q, t)) \times f(p, q, t) = \langle 0, 0, 0 \rangle,$$

esto es, $f(p, q, t)$ es paralelo a $s(p, t) - s(q, t)$, lo que significa que la línea de acción de $f(p, q, t)$ está a lo largo de la línea que conecta ambas partículas. Un argumento similar demuestra que $f(q, p, t)$ es paralelo a $s(p, t) - s(q, t)$. Una consecuencia de este resultado es que toda fuerza interna es una fuerza de atracción o de repulsión.

El axioma final, P7, formula la segunda ley del movimiento de Newton, que originalmente enunció como sigue:[39] El cambio de movimiento es proporcional a la fuerza motriz impresa, y se efectúa en la dirección de la línea recta en que se imprime esa fuerza. Algunos autores[40] han propuesto que se convierta la segunda ley, esto es, P7, en una definición de la fuerza total que actúa sobre un punto material. Parece haber dos serias objeciones contra este procedimiento. Prohíbe dentro del marco axiomático cualquier análisis de las fuerzas internas y externas que actúan sobre un punto material. Esto es, si todas las nociones de fuerza se eliminan como primitivas y P7 se usa como definición, entonces las nociones de fuerza interna y externa no son definibles dentro del marco axiomático dado. En segundo lugar, convierte en una definición una de las más importantes leyes de la mecánica, que es algo que los convencionalistas encuentran atractivo, pero que difícilmente concuerda con una filosofía de la ciencia de orientación empírica. Con relación a la axiomatización dada aquí, puede demostrarse por el principio de Padoa que las nociones de masa y de fuerza interna son cada una independientes de las restantes nociones primitivas. (Dejamos esta demos-

[39] Obra citada, Pág. 13.

[40] Por ejemplo, G. Kirchhoff, en su clásica obra, *Vorlesungen über mathematische Physik*, I: Mechanik, 1878.

tración como ejercicio). Por otra parte, la noción de fuerza externa (resultante) es definible, puesto que la ecuación del Axioma P7 puede escribirse:

$$g(p, t) = m(p)D^2s_p(t) - \sum_{q \in P} f(p, q, t).$$

Sin embargo, si la noción de fuerza externa dada se hubiera tomado como primitiva, podría haberse fácilmente demostrado que era independiente. Se usó aquí la resultante de las fuerzas externas dadas para evitar la complicación de introducir una serie infinita (correspondiente a un número infinito de fuerzas externas dadas).

Los lectores familiarizados con los escritos técnicos sobre los fundamentos de la mecánica pueden encontrar extraña la pretensión de que las nociones de masa y de fuerza interna son independientes de las otras nociones primitivas, puesto que se discuten tan ampliamente las definiciones de estas nociones. Sin embargo, estas discutidas definiciones, como la "definición" de Mach de la masa relativa de dos cuerpos como la razón inversa de sus aceleraciones "mutuamente inducidas" cuando están aislados de otros cuerpos,[41] son sólo definiciones en cierto ingeuno sentido que tiene poca conexión con la teoría de la definición sostenida en el Cap. 8.

Pasamos ahora a algunos teoremas que iluminan varias facetas de la mecánica de partículas. El primer teorema se relaciona con el centro de masa de un sistema de la mecánica de partículas (formalmente definidos adelante) y que se formulan, generalmente, en el modo subjuntivo: el centro de masa de un sistema se mueve *como si* toda la masa *estuviera allí concentrada* y la resultante de todas las fuerzas actuara allí. En años recientes, algunos filósofos han proclamado que el uso de condicionales contrarias a los hecos y, en consecuencia, el uso del modo subjuntivo no aleticofuncional son necesarios para formular adecuadamente leyes empíricas (aunque esta necesidad no se proclama para el ejemplo dado aquí). Sin embargo, no conozco ningún ejemplo sistemático como el presente que no pueda ser reformulado satisfactoriamente en el modo indicativo y ciertamente no conozco ningún teorema en rama alguna de la ciencia empírica cuya demostración estribe en alguna peculiar propiedad no aleticofuncional de condicionales contrarias a los hechos.[42] Una definición formal elemental de la función centro de masa. es directa.

[41] E. Mach, *Science of Mechanics*, 5a. Ed. americana, La Salle, Ill., 1942, Págs 264 a 277. Desde un punto de vista formal el libro de Mach es una masa de confusiones.

[42] Esto no significa que se niegue que la lógica del modo subjuntivo es de considerable interés filosófico general.

DEFINICIÓN E2. *Si $t \in T$, entonces*

$$c(t) = \frac{\sum_{p \in P} m(p)s(p,t)}{\sum_{p \in P} m(p)}.$$

Y el teorema apropiado es:

TEOREMA E1. Si $t \in T$, entonces

$$\sum_{p \in P} m(p)D^2c(t) = \sum_{p \in P} g(p,t).$$

Nótese que el movimiento depende sólo de las fuerzas externas, puesto que las internas se cancelan mutuamente; de esta observación, así como del Axioma P7 y de la definición anterior, inferimos inmediatamente el teorema.

En el apartado número tres de este capítulo se discutió con cierto detalle la noción de que dos modelos de una teoría son isomórficos y se hizo notar que una de las más interesantes aplicaciones de la idea de isomorfismo, es para establecer un teorema de representación de la teoría. Algunas veces no es posible demostrar un interesante teorema de representación. Cuando se da esta situación, es natural preguntar si puede demostrarse para la teoría un teorema de *incrustación*, esto es, demostrar que hay una interesante clase K de modelos tales, que todo modelo de la teoría es isomórfico con respecto de un submodelo de un modelo que pertenece a K. La definición exacta de un submodelo varía de una teoría a otra. Si, por ejemplo $\mathfrak{A} = \langle A, \circ \rangle$ es una álgebra, esto es, A es un conjunto y $\circ$ es una operación binaria de $A \times A$ a A, entonces una álgebra $\langle A' \circ' \rangle$ es una *subálgebra* de $\mathfrak{A}$ si A' es un subconjunto de A y $\circ' = \circ \cap (A' \times A' \times A')$, esto es, $\circ'$es la operación $\circ$ restringida a A'. En el caso de la teoría de la mecánica de puntos materiales, la relativa complejidad de las nociones primitivas trae como consecuencia que sean naturales varias definiciones alternativas de submodelos. Por ejemplo, en un subsistema (esto es, submodelo) de un sistema de mecánica de puntos materiales, ¿permitiríamos que el subsistema tuviera un intervalo de tiempo más pequeño? La decisión un tanto arbitraria hecha aquí es para mantener igual el intervalo de tiempo en el subsistema. La parte decisiva de la definición de subsistemas peculiar a la mecánica es el manejo de fuerzas internas. Las fuerzas internas que actúan sobre un punto material, debidas a partículas que no están en el subsistema, se suman a la fuerza externa resultante ejercida sobre el

punto material, procedimiento que parece concordar con la idea intuitiva que la física tiene de un subsistema.

DEFINICIÓN 3. *Sea* $\mathfrak{P} = \langle P, T, s, m, f, g \rangle$ *un sistema de mecánica de puntos materiales; sea* P′ *un subconjunto no vacío de* P; *sean* s′ *y* m′ *las funciones* s *y* m, *con sus primeros argumentos restringidos a* P′; *sea* f′ *la función* f, *con sus primeros dos argumentos restringidos a* P′, *y sea* g′ *una función tal que para toda* p *de* P′ *y* t *de* T

$$g'(p, t) = g(p, t) + \sum_{q \in P \sim P'} f(p, q, t).$$

Entonces $\langle P', T, s', m', f', g' \rangle$ *es un subsistema de* $\mathfrak{P}$.

Esta definición podía haber sido puesta en un lenguaje más sinforemático. Podríamos haber escrito:

$$P' \subseteq P$$
$$P' \neq \Lambda,$$
$$s' = s|(P' \times T)$$
$$m' = m|P'$$
$$f' = f|(P' \times P' \times T),$$

donde, en general, si f es una función, $f|A$ es la función h, tal que

$$h = f \cap (A \times \mathscr{R}(f)),$$

esto es, $f|A$ es la función f con su dominio restringido a A.

Dejamos como ejercicio la demostración del

TEOREMA 2. *Todo subsistema de un sistema de mecánica de puntos materiales es él mismo un sistema de mecánica de puntos materiales.*

Por supuesto, el Teorema 2 no constituye un teorema de incrustación y es éste el problema al cual pasamos. Una clase filosófica y físicamente interesante de sistema de mecánica de puntos materiales es la de sistemas *aislados*, esto es, aquellos para los cuales es igual a cero la fuerza externa resultante ejercida sobre cada partícula (más precisamente, es igual al vector cero $\langle 0, 0, 0 \rangle$).

DEFINICIÓN 4. *Un sistema de mecánica de puntos materiales* $\mathfrak{P} = \langle P, T, s, m, f, g \rangle$ *es aislado si y sólo si para toda* p *de* P *y toda* t *de* T

$$g(p, t) = \langle 0, 0, 0 \rangle.$$

La noción de sistema aislado está relacionada con la noción del físico de un sistema mecánico *cerrado* o *independiente* que, hablando en términos generales, es un sistema aislado, en que la fuerza interna de una partícula cualquiera sobre otra se puede expresar como función solamente de la distancia entre las dos partículas. La energía mecánica total de los sitemas cerrados es constante en el tiempo; más importante, cuando las fuerzas internas no dependen explícitamente del tiempo podemos considerar el análisis causal de los movimientos de las partículas como completo en cierto sentido. Desgraciadamente, el problema de incrustar cualquier sistema de mecánica en un sistema cerrado es demasiado complicado para discutirlo aquí y nos conformaremos con un resultado más débil de incrustación, relativo a sistemas aislados. Primeramente, definimos la noción de dos sistemas de mecánica *equivalentes;* la idea intuitiva es que los dos sistemas son idénticos, excepto por lo que respecta a cada una de las fuerzas que actúan sobre los puntos materiales; la fuerza resultante sobre un punto material es la misma en ambos sistemas.

DEFINICIÓN 5. *Dos sistemas de mecánica de partículas* $\mathfrak{P} = \langle P, T, s, m, f, g \rangle$ *y* $\mathfrak{P}' = \langle P', T', s', m', f', g' \rangle$ *son equivalentes si y sólo si*

$$P = P'$$
$$T = T'$$
$$s = s'$$
$$m = m'.$$

Nótese que esta noción de equivalencia es uno de varios conceptos que son al mismo tiempo más débiles y más fuertes que la noción natural de isomorfismo;[43] es más débil, en que dos sistemas equivalentes no tienen la misma estructura de fuerzas individuales; pero es más fuerte en que dos sistemas equivalentes deben ser cinemáticamente idénticos y también idénticos en sus funciones de masas.

Dejamos como ejercicio la demostración de los dos teoremas siguientes; justifican el uso del término "equivalencia" en la Definición 5.

TEOREMA 3. *La relación de equivalencia entre sistemas de mecánica de partículas es reflexiva, simétrica y transitiva en el conjunto de todos los sistemas de mecánica de partículas.*

[43] Cual sea exactamente *la* definición natural de isomorfismos para sistemas de mecánica de puntos materiales es una cuestión un tanto complicada que no trataremos de discutir aquí.

El teorema siguiente formula con precisión la idea de que las fuerzas resultantes son idénticas en dos sistemas equivalentes.

TEOREMA 4. *Si* $\mathfrak{P} = \langle P, T, m, s, f, g\rangle$ *y* $\mathfrak{P}' = \langle P', T', m', s', f' g'\rangle$ *son dos sitemas equivalentes de mecánica de partículas, entonces para toda p de P y toda t de T*

$$\sum_{q\in P} f(p, q, t) + g(p, t) = \sum_{q\in P'} f'(p, q, t) + g'(p, t).$$

(Nótese que puesto que $\mathfrak{P}$ y $\mathfrak{P}'$ son equivalentes, ambos miembros de la ecuación están bien definidos.)

El teorema de incrustación que deseamos ahora demostrar, es que todo sistema de mecánica de partículas es equivalente a un subsistema de un sistema aislado. Este teorema está estrechamente relacionado con algunas posiciones relativas a los fundamentos de la física, históricamente famosas. Por ejemplo, Roger Joseph Boscovich, el prominente físico jesuita del siglo XVIII sostuvo que la materia del universo está compuesta de un número finito de puntos inextensos y que las únicas fuerzas del universo son fuerzas de atracción y de repulsión que actúan entre los puntos y que satisfacen nuestros Axiomas P5 y P6. Todos los fenómenos observados de la naturaleza han de ser explicados solamente en términos de la distribución y movimiento de estos puntos y de las fuerzas existentes entre ellos. Ciertamente, Boscovich[44] consideraría nuestro teorema de incrustación como un paso definitivo hacia el establecimiento de su tesis relativa a la naturaleza del universo. Naturalmente no estamos aquí sentando esas pretensiones, pues no deseamos aseverar que las nuevas partículas, sumadas para formar el sistema aislado en que está incrustado un sistema dado, tengan una existencia física real.

Nuestro teorema de incrustación está más cercano al análisis de los fundamentos de la mecánica suministrados por el gran físico del siglo diecinueve Heinrich Hertz, cuyos *Principios de Mecánica* fueron publicados por primera vez en Alemania en 1894.[45] La idea básica de su sistema es que cualquier sistema mecánico complicado ha de explicarse suponiendo que es parte de un sistema más grande, la posición de cuyas partículas explica el carácter del sistema original en forma natural y sencilla. Es de observarse que Hertz solamente consideraría la incrustación de un sistema en un sistema cerrado, no solamente aislado, como un resultado totalmente satisfactorio. Pasamos ahora a nuestro teorema.

---

[44] El principal tratado de Boscovich es su *Theoria Philosophiae Naturalis*, publicado por primera vez en Viena en 1758.

[45] Traducción inglesa reimpresa en 1956 por Dover Publications, Nueva York.

TEOREMA 5. *Todo sistema de mecánica de puntos materiales es equivalente a un subsistema de un sistema aislado de mecánica de puntos materiales.*[46]

DEMOSTRACIÓN. Para simplificar, consideraremos sólo un sistema bidimensional, esto es, un sistema para el cual es igual a cero la tercera componente de todos los vectores de posición y de fuerza. Sea $\mathfrak{P} = \langle P, T, s, m, f, g\rangle$ este sistema arbitrario de mecánica. La idea de la demostración consiste en introducir para cada p de P tres nuevas partículas, tales que las cuatro partículas juntas están colocadas simétricamente con respecto de los dos ejes de coordenadas. A continuación, se descompone la fuerza total resultante sobre p en fuerzas internas entre la partícula dada p y dos de las nuevas partículas. La partícula nueva restante se necesita para suministrar un equilibrio apropiado de fuerzas internas sobre las dos primeras partículas nuevas.

Podemos ahora pasar a la demostración formal. Es conveniente introducir dos funciones que particularicen las primeras dos componentes de un vector. Si $\langle x_1, x_2, x_3\rangle$ es un vector, entonces

$$\varphi_1(\langle x_1, x_2, x_3\rangle) = x_1,$$
$$\varphi_2(\langle x_1, x_2, x_3\rangle) = x_2.$$

Para simplificar la notación, y sin pérdida esencial de generalidad, podemos suponer que P tiene exactamente dos partículas, digamos,

$$P = \{p, q\}.$$

Usamos ahora el sistema $\mathfrak{P} = \langle P, T, s, m, f, g\rangle$ para definir un nuevo sistema consistente en seis partículas diferentes:

$$P' = \{p, q, p_1, p_2, p_3, q_1, q_2, q_3\}$$
$$T' = T$$
$$m'(p_i) = m'(p) = m(p) \qquad \text{para } i = 1, 2, 3$$
$$m'(q_i) = m'(q) = m(q) \qquad \text{para } i = 1, 2, 3,$$

[46] Este teorema está estrechamente relacionado con el Teorema 8 del Trabajo de McKinsey, Sugar y Suppes anteriormente citado.

y para toda t de T'

$$
\begin{aligned}
s'(p, t) &= s(p, t)\\
s'(p_1, t) &= \langle -\varphi_1[s(p, t)], \varphi_2[s(p, t)], 0\rangle\\
s'(p_2, t) &= \langle \varphi_1[s(p, t)], -\varphi_2[s(p, t)], 0\rangle\\
s'(p_3, t) &= \langle -\varphi_1[s(p, t)], -\varphi_2[s(p, t)], 0\rangle\\
s'(q, t) &= s(q, t)\\
s'(q_1, t) &= \langle -\varphi_1[s(q, t)], \varphi_2[s(q, t)], 0\rangle\\
s'(q_2, t) &= \langle \varphi_1[s(q, t)], -\varphi_2[s(q, t)], 0\rangle\\
s'(q_3, t) &= \langle -\varphi_1[s(q, t)], -\varphi_2[s(q, t)], 0\rangle
\end{aligned}
$$

y

$$f'(p, q, t) = f'(q, p, t) = \langle 0, 0, 0\rangle.$$

Sea

$$
\begin{aligned}
F(p, t) &= g(p, t) + f(p, q, t)\\
F(q, t) &= g(q, t) + f(q, p, t),
\end{aligned}
$$

esto es, $F(p, t)$ es la fuerza total resultante ejercida sobre la partícula p en el instante t y en el sistema $\mathfrak{P}$. Entonces, hacemos:

$$
\begin{aligned}
f'(p, p_1, t) &= \langle \varphi_1[F(p, t)], 0, 0\rangle\\
f'(p_1, p, t) &= -f'(p, p_1, t)\\
f'(p, p_2, t) &= \langle 0, \varphi_2[F(p, t)], 0\rangle\\
f'(p_2, p, t) &= -f'(p, p_2, t)\\
f'(p, p_3, t) &= f'(p_3, p, t) = \langle 0, 0, 0\rangle\\
f'(p_1, p_2, t) &= f'(p_2, p_1, t) = \langle 0, 0, 0\rangle\\
f'(p_1, p_3, t) &= \langle 0, \varphi_2[F(p, t)], 0\rangle\\
f'(p_3, p_1, t) &= -f'(p_1, p_3, t)\\
f'(p_2, p_3, t) &= \langle \varphi_1[F(p, t)], 0, 0\rangle\\
f'(p_3, p_2, t) &= -f'(p_2, p_3, t).
\end{aligned}
$$

Reemplazando 'p' por 'q', 'p_1' por 'q_1', etc., en las diez líneas precedentes obtenemos la nueva función de la fuerza, f', para las partículas q, q_1, q_2, q_3. (No es necesario asentar estas diez líneas adicionales). Finalmente, definimos la función g' de la fuerza externa, de manera que para todas las partículas de P' y para todas las t de T' la fuerza externa es el vector nulo $\langle 0, 0, 0\rangle$.

Para completar la demostración del teorema necesitan ser verificadas dos cosas (que dejamos como ejercicio):

(i) El sistema $\mathfrak{P}' = \langle P', T', s', m', f' g' \rangle$ es un sistema de mecánica de puntos materiales, esto es, satisface los Axiomas P1 a P7 (si $\mathfrak{P}'$ es un sistema de mecánica de puntos materiales, se sigue inmediatamente de la definición de g' que está aislado).

(ii) $\mathfrak{P}$ Es equivalente a un subsistema de $\mathfrak{P}'$ Q.E.D.

La teoría de la mecánica es como la teoría de la probabilidad en que muchos de los más interesantes y difíciles problemas se relacionan con clases especiales de modelos de la teoría. Probablemente el ejemplo más famoso en mecánica es el problema de tres cuerpos, cuando las fuerzas que actúan sobre ellos son las fuerzas gravitacionales de atracción mutua, que varían en razón inversa del cuadrado de las distancias. La solución de Newton al problema de dos cuerpos constituyó una deducción de las tres leyes cinemáticas de Kepler de la ley dinámica de la gravitación. Desgraciadamente, los confines del presente apartado son demasiado estrechos para permitir un ulterior desarrollo sistemático de la mecánica. Cierta cantidad de material adicional se incluye en los ejercicios.

EJERCICIOS

1. Usando las definiciones dadas al comienzo de este apartado y los hechos conocidos sobre los números reales, tales como los axiomas y teoremas del Cap. 7, demuéstrese:

(a) La suma vectorial es conmutativa y asociativa.

(b) La multiplicación de un vector por un número real es distributiva con respecto de la adición vectorial, esto es, si a es un número real y, x y y son vectores, entonces

$$a(x + y) = ax + ay.$$

(c) La multiplicación de un vector por un número real es asociativa, esto es, si a y b son números reales y x es un vector, entonces

$$(a + b)x = ax + bx.$$

(d) La multiplicación de un vector por un número real es asociativa, si a y b son números reales y x es un vector, entonces

$$a(bx) = (ab)x.$$

(e) La operación de multiplicación escalar es conmutativa.

(f) La operación de multiplicación escalar es distributiva con respecto de la adición vectorial, esto es, si x, y y z son vectores, entonces

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

(g) La operación de multiplicación vectorial es distributiva con respecto de la suma vectorial, esto es, si x, y y z son vectores, entonces

$$x \times (y + z) = x \times y + x \times z.$$

(h) La operación de multiplicación vectorial es anticonmutativa, esto es, si x, y z son vectores, entonces

$$x \times y = -y \times x.$$

(i) (DESIGUALDAD DE SCHWARZ) Si x y y son vectores, entonces

$$|x \cdot y| \leq |x|\,|y|.$$

2. El conjunto de todos los vectores es un grupo con respecto a cuál de las siguientes operaciones (si lo es, demuéstrese, si no, dése un contraejemplo):

(a) Suma vectorial.
(b) Sustracción vectorial.
(c) Multiplicación escalar.
(d) Multiplicación vectorial.

3. Vuelva a formular la Definición 1 en el estilo de la Definición A del § 12.2
4. Demuéstrese el Teorema 2.
5. Demuéstrese el Teorema 3.
6. Demuéstrese el Teorema 4.
7. Complétese la demostración del Teorema 5.
8. Para sistemas tridimensionales de mecánica, ¿cuántas partículas nuevas deben añadirse cuando se incrusta el sistema en un sistema aislado? Demuéstrese el Teorema 5 para sistemas generales tridimensionales.
9. Usando el principio de Padoa demuestra que cada una de las nociones de masa y de fuerza interna son independientes de las otras nociones primitivas de la mecánica de puntos materiales.
10. Demuéstrese por el método de interpretación que cada uno de los axiomas dinámicos de la Definición 1 son independientes de los axiomas restantes.
11. Defina el momento de una fuerza con respecto de un punto fijo (puntos y vectores son las mismas entidades, a saber, triples ordenadas de números reales) y demuestre que el momento total de las fuerzas internas de un sistema de mecánica de puntos materiales, es el vector nulo.[47]
12. Defina el momento angular de un sistema de puntos materiales con respecto de un punto fijo y demuestre que la velocidad de cambio del momento angular de un sistema, con respecto de un punto, es igual al momento total de las fuerzas externas con respecto de ese punto.
13. Formule exactamente y deduzca de los axiomas de la Definición 1 una forma de las ecuaciones de Lagrange para sistemas de mecánica de puntos materiales.
14. Deduzca las tres leyes del movimiento de Kepler de la hipótesis de Newton relativa a las fuerzas gravitacionales.

[47] Este y los ejercicios restantes requieren cierto conocimiento intuitivo de la mecánica de puntos materiales, pero las definiciones usuales deben ser reformuladas como definiciones elementales que satisfagan las reglas del Cap. 8.

INDICE

— A —

Adams, E., 332
Aleticofuncional, 27, 30
Algebra booleana, 255, 257
Algebras, 310, 316
  isomórficas, 320
  subálgebras, 365
Argumento inválido, 58, 103
Argumento válido, 98, 156
  y tablas de verdad, 49
Aristóteles, 360
Axioma de oclusión, 316
Axiomas, 99
  de oclusión, 316
  independencia de los, 106
  para álgebras booleanas, 255
  para espacios de probabilidad, 340
  para grupos, 146, 310, 316, 317
  para la mecánica de puntos materiales, 359
  para medición de masa, 136
  para números aritméticos o reales, 162
  para preferencia, 217
Axiomatización; 305 y siguientes, *véase también* Axiomas
  en lógica de primer orden, 306
  en teoría de los conjuntos, 307

— B —

Bayes, teorema de, 346
Bernays, P., 98, 148*n*
Bernoulli, J., 350
Bicondicional, 33
Birkhoff, G., 324
Boscovich, R. J., 368 y sig.
Bourbaki, N., 308n
Bush, R. R., 346

— C —

Campos de conjuntos, 339
Carnap, R., 213, 355n
Casos en la demostración, 194
Cantor, G., 299n
Cinemática, 358, 361
Clase, 225, *véase también* Conjuntos
Clase de equivalencia, 271
Comillas, 163
Complemento de conjuntos, 237
Composición de funciones, 293
Concatenación, 169
Condición, necesaria, 32, 191
  suficiente, 32, 191
Condicional contraria a los hechos, 364
Conectivos oracionales, 25
  aleticofuncionales, 27, 30
Conjunción, 22, 34
Conjunto, vacío, 226, 233
Conjuntos, 225
  complemento, 237
  diferencia, 234
  extensionalidad, 226
  inclusión, 230
  intersección, 234
  mutuamente exclusivos, 234
  unión, 234
Consecuencia lógica, 47, 102, *véase también* Inferencia
Consistencia, 65, 103
Constante individual, 102, *véase también* Nombre propio,
  definición, de, 206
  independencia de la, 218
Contextos no extensionales, 143
Contradicción, 65
Contraejemplo, 104

Convencionalismo, 360
Copi I, 117
C. P., 55
Criterio, de eliminabilidad, 200
  de no-creatividad, 200
Cuantificador, 78
  alcance de un, 84
  existencia, 79
  intercambio de, 124
  reglas que rigen el, 157
  universal, 79, 92, 93
  y citación, 167
Cuasiordenación, 273

— CH —

Church, A., 104, 297

— D —

Davidson, D., 329n
Deducción, *véase también* Inferencia
Demostración oracional, 55
Deducción falaz, *véase* Inferencia falaz
Definición, 197
  como identidad, 208
  condicional, 212
  criterio de eliminabilidad, 200
  de constantes individuales, 206
  de la división, 210, 213
  de símbolos de operación, 205, 209
  de símbolos de relación, 202
  criterio de no-creatividad, 200
  elemental, 329
  y axiomatización, 307 y sig.
Definición recursiva, de fórmulas, 83
Definiendo, 202
Definiente, 202
De Morgan, A., 131
Demostración, casos alternativos en la, 195
  condicional, 55
  estrategia de la, 57, 91, 181
  falaz, 182
  falta de plenitud de la, 105
  indirecta, 67
  informal, 171
Derivada, 358
Dewey, J., 322n
Diádico, *véase* Dinario
Diagramas, de Hasse, 273
  de Venn, 245 y sig.
Diferencia de conjuntos, 234
Diferenciable, 358
Dinámica, 359
Disyunción, 28, 34
Dominio, de individuos, 98, y sig. 236
  de interpretación, 106

— E —

EG, 118
Elementos inverso, 146, 174
Entidad lingüística, 287
  sinforemática, 287
Epistemología, 123
Epsilon, 226
Equivalencia, 33, 35
  tautológica, 40, 61, 125
ES, 118
Espacio de muestra, 340
Especificación existencial, 116
Especificación universal, 92
  restricción a la, 128
Estes, W. K., 346
Estructuras, de diferencias, 327
  simples de relación, 310
Euclides, 303
Euler, diagramas de, 245 y sig.
Evento, acontecimiento o suceso, 336, y sig.
  atómico, 351
Extensionalidad, principio de, 143
  conjuntos, 226

— F —

Falacia, de afirmación del consecuente, 50
  de negar el antecedente, 50
Falta de plenitud, 105
Feller, W., 354n
Forma lógica, 102
Fórmula, 83
  abierta, 144
  atómica, 83
  consistencia, 103
  definición recursiva de, 83
  equivalencia lógica de una, 160
  universalmente válida, 102
Frege, G., 20
Fuerza, 359
Función, 283 y sig.
  alcance de una, 287

composición de, 293
constante, 299
definición, 283 y sig.
dominio de una, 287
inversa, 290
mapa de una, 288
oracional, 170
principal inversa, 292
proposicional, 170
Función uno-uno, 290

— G —

Generalización existencial, 118
restricción a la, 127, 130, 131
Generalización universal, 93
restricción a la, 128
Geometría euclidiana, 303 y sig.
Gödel, K., 105
Gramática, 73
Granville, W. A., 283n
Grupo de transformaciones, 323

— H —

Hamel, G., 362n
Hasse, H., 273n
Hertz, H., 368
Hilbert, D., 148n, 305
Huntington, E. V., 97, 221, 255n

— I —

Idempotencia, 255
Identidad, 141
elemento de, 146, 173, 255
interpretación de la, 148
Imagen, 288
Implicación, 29, 35
tautológica, 39, 54, 61
Imprecisión, 94, 117, 131
Incidencia de la variable, libre, 85
ligada, 85
Inclusión de conjuntos, 230
Independencia de axiomas, 106
Independencia de símbolos primitivos, 217
Inducción matemática, 334n
Inferencia, criterios de la, 46
falaz, 49, 127, 138-139, 182 y sig.
fórmulas lógicamente equivalentes, 160
legitimidad de la, 49
plenitud de la, 50
regla reducida de, 68, 155
reglas de identidad, 144
resumen de las reglas de, 137
teoría oracional de la, 46
Instanciación universal, 93
Interpretación, aritmética, 99
de nombres ambiguos, 121
de nombres propios, 100, 111
de predicados, 109
de símbolos de operación, 101, 113
de variables libres, 99, 111
definición de, 102
dominio de individuos para, 98
oracional, 46
reglas de, 106
verdadera, 103
y argumento inválido, 58
Intersección de conjuntos, 234
Intervalo, abierto, 356
cerrado, 356
de números reales, 356
Isomorfismo, 320
de álgebras, 320
de estructuras de diferencias, 327
de estructuras simples de relación, 321
y categoricidad, 322
Jeffries, H., 355n

— K —

Kepler, J., 371
Kirchhoff, G., 363n
Kleene, S. C., 105
Kolmogorov, A. N., 338n

— L —

Laplace, P., 351n
Leibniz, G., 19
ley de, 143
Lesniewski, S., 200
Ley, de absorción, 255
de exportación, 61
de identidad, 143
de importancia, 61
del absurdo, 61
del tercero excluido, 62, 255
de la Contraposición, 61, 255
de la Separación, 59
distributiva, 173, 175, 255

Ley asociativa, para grupos, 146
para la edición, 173
para la multiplicación, 173
Ley conmutativa, para la adición, 173
para intersección, 255
para la multiplicación, 173
para la unión, 255
Ley de la cancelación, para grupos, 146
para la multiplicación, 194
Leyes de Morgan, 61, 244, 255
Lógica, con identidad, 148
de predicados de primer orden, 104
definición de, 22
teorema de, 149
Longley, W. R., 283n
Luce, R. D., 282

— M —

Mach, E., 221, 364
MacLane, S., 324
Mapa, 288
Masa, 359
centro de, 364
definición de, 364
medición de, 135
Matriz, 114
Maxwell, J. C., 361
McKinsey, J. C. C., 219n, 359n, 369n
McNaughton, R., 248n
Mecánica; 356 y siguientes, *véase también* Sistema de mecánica de puntos materiales
Medición de masas, 135
de utilidad, 327
Medida de probabilidad enumerablemente aditiva, 340
Medida de probabilidad finitamente aditiva, 340
Medida de probabilidad finitamente aditiva, 340
Membrecía, 225
Mención, 163
Matemáticas, 311
Método axiomático, 303 y siguientes
Mises R., von, 354n
Modelo, *véase* Interpretación y Categoricidad, 322
Modelo para una teoría, 311
Modo subjuntivo, 364
Modus ponendo ponens, 60
Modus tollendo ponens, 60
Modus tollendo tollens, 60
Mosteller, F., 346
Muestreo al azar, 350
Mutuamente exclusivo, 234

— N —

Nagel, E., 354n
Negación, 25, 34
Negación del antecedente, falacia de, 50
Newton, I., 363, 371
Nociones primitivas, 315, *véase también* Símbolos primitivos
Nombres, 167
véase también Nombres ambiguos, Nombres propios
Nombres ambiguos, 117
interpretación, 121
restricciones al uso de, 119, 151
subíndices en, 127
Nombres propios, inferencias, 84, 95
véase también Constantes individuales
Notación lambda, 298
No, *véase* Negación
Número, racional, 174
real, 174

— O —

"O", 28 y siguientes
Operación binaria, 288
y relación, 288
Operación negativa, 188
Operaciones,
en conjuntos, 233 y siguientes, 252 y siguientes
con funciones, 289 y siguientes
con relaciones, 278 y siguientes
Operador de abstracción, 298
de descripción, 208
Oración atómica, 35
bicondicional, 29
condicional, 29
de reducción, 213
definición de, 86
véase también Fórmula
Ordenación, estricta simple, 275
parcial, 273
parcial estricta, 273
simple, 274

— P —

P, 54
Padoa, A., 217, 218
Paréntesis, 33
Par ordenado, 259
Peano, G., 208
Peirce, G., 39n
Pieri, M., 305
Plenitud de las leyes de inferencia, 50, 105
Poincaré, H., 360n
Posibilidad, 41
Postulado, *véase* Axiomas
Potencia o relación, 333-334
Putnam, H., 124
Predicado, colectivo, 241
de una posición, 77
distributivo, 241
interpretación de, 109, 110
sinforemático, 307 y siguientes
véase también Símbolo de relación, 76
Predicado sinforemático, 287
Predicción, 19
Preferencia, *véase* Teoría de la preferencia,
Premisas, 45, 54
consistencia de las, 64
Principal inverso, 292
Principio, de indiferencia, 350
de Padoa, 217
Probabilidad, 335
a posteriori, 347-348
condicional, 343
distribución de, 342
espacio de, 340
inversa, 348
inversa máxima, 354
medida de la, 340
probabilidad a priori, 348
términos de, 338
Procedimiento de decisión, 104
Producto, cartesiano de conjuntos, 261
de cruz de los conjuntos, 261
de cruz de vectores, 357
escalar de dos vectores, 357
interior de dos vectores, 357
relativo, 279
vectorial, 357
Propiedad de oclusión, 113
Prueba, de consistencia, 66
de inconsistencia, 63
de independencia de axiomas, 106
de la independencia de símbolos, 217
Pseudo operación, 198, 205

— Q —

Q.E.D., 176
Quine, W. V., 117, 166n

— R —

R.A.A., 68
Reduccionismo, 213n
Reductio ad absurdum, 67
Regla derivada de inferencia, 68, 155
Reglas efectivas, 155
Reichenback, H., 355n
Relación, 261 y siguientes
antisimétrica, 266
asimétrica, 265
campo de, 263
contradominio de, 263
binaria, 262
conectada, 268
de equivalencia, 270
definición de, 262
dominio de, 263
fuertemente conectada, 268
intransitiva, 268
irreflexiva, 265
muchos-uno, 288
recíproco de, 279
reflexiva, 264
símbolo de, 101, *véase también* Predicado
simétrica, 265
subrelación, 279
ternaria, 262, 288
transistiva, 267
universal, 278
Relaciones, ordenación de, 272 y sig.
operaciones con, 278 y sig.
propiedades de las, 264 y sig.
Representación para grupos, 234
teorema de la, 323
y medición, 326 y sig.
y reduccionismo, 332
Rubin, H., 359n
Russell, B., 20, 97, 98n

— S —

Savage, L. J., 355n
Secuencia, 240n
finita, 261
Segunda ley del movimiento de Newton, 363
Semántica, 311n
Semi-orden, 282
Si. . .entonces, *véase* Implicación
Significado, 103, 168, 214
Silogismo, 250 y sig.
Símbolo de operación, 205
definición de, 213
independencia del, 218
Símbolo de relación, definición de, 202
independencia del, 218
Símbolos primitivos, 314, *véase también* Nociones primitivas
independencia de los, 217
Simons, L., 124
Sintaxis lógica, 311n
Sistema de mecánica de puntos materiales aislado, 366
definición, 359
equivalente, 367
subsistema, 366
Smith P. F., 283n
Sólo si, 31
Subíndices, *véase* Nombres ambiguos
Subrelación, 279
Sugar, A. C., 359n, 369n

— T —

T, 54
Tablas de verdad, 34
Tarski, A, 55n, 105n, 169n, 172n, 219n
Tautología, 36 y siguientes, 61
definición de, 38
e interpretación oracional, 47
Teorema, de deducción, 55
de incrustación, 365
categórica, 322
de la preferencia, 97, 106, 116, 217, 282, 318
Teoría de los conjuntos 225
axiomática, 307-308
Teoría de los grupos, 146
axioma de representación de la, 323
axiomas de la, 146, 154, 310, 316, 317
Tercera ley del movimiento de Newton, 362
Términos, 74
Términos indefinidos, *véase* Símbolos primitivos
triádico, *véase* Ternario

— U —

UG, 94
Unión de conjuntos, 234
Universo del discurso, *véase* Dominio de individuos
US, 93
Uso y mención, 163

— V —

Validez, 49, 98, *véase también* Validez universal
Validez lógica, *véase* Validez universal, Validez
Validez universal, 102, 156
y teorema de lógica, 149
Variable, 74 y sig.
imprecisa, 94
incidencia libre de una, 85
numérica, 166
oracional, 165
Variables libres, 85, 184
interpretación de las, 108, 113
Variables ligadas, 85
incidencia de, 85
interpretación de la, 107
Veblen, O., 322n
Vector, 356
valor absoluto de un, 357
Vel, 28
Venn, J., 245n
Verdad, e interpretaciones, 103
valores de, 38
y lógica, 269
y teoremas de lógica, 149

— W —

Whitehead, A. N., 97, 98
Wittgenstein, L., 39

— Y —

"Y", 25 y sigts.

NOTAS

NOTAS

NOTAS

NOTAS

NOTAS

ESTA IMPRESION DE 1,000 EJEMPLARES
SE TERMINO EL 30 DE SEPTIEMBRE DE
1970, EN LOS TALLERES DE LA COMPAÑIA
EDITORIAL CONTINENTAL, S. A., MEXICO

INTRODUCCION A LA LOGICA SIMBOLICA

Por

PATRICK SUPPES

Introducción a la Lógica Simbólica es una presentación vigorosa y directa de la lógica moderna y de la teoría elemental de los conjuntos, organizada con habilidad y criterio. El autor deriva un sistema original de lógica de primer orden que se aproxima a la escritura informal de pruebas de matemáticas y otras ciencias.

PRINCIPIOS DE INFERENCIA Y DEFINICIÓN

PARTE I

En los primeros cinco capítulos, el profesor Suppes desarrolla un sistema de deducción natural, incluyendo la teoría de la identidad. Este material es seguido por capítulos sobre la transición de las pruebas informales a las formales y sobre la teoría de la definición. Estos capítulos, creemos, son el primer tratamiento satisfactorio de estos tópicos que aparecen en un libro de consulta.

PARTE II

TEORÍA ELEMENTAL DE LOS CONJUNTOS

Los primeros tres capítulos están dedicados a los conjuntos, propiedades de orden de relaciones y da una exacta definición de funciones de los conjuntos teóricos. El capítulo final explora los fundamentos teóricos de los conjuntos del método axiomático, incluyendo desarrollos axiomáticos de tópicos que van desde las ciencias sociales y de la física, lo mismo que desde las matemáticas.

Este nuevo estudio está escrito en forma vívida y precisa y constituye un magnífico texto para los cursos sobre la lógica moderna.